Measurement of Spinal Stiffness in Predicting Treatment Response in Low Back Pain Patients

by

Maliheh Hadizadeh Bajestani

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Rehabilitation Science

Faculty of Rehabilitation Medicine

University of Alberta

ABSTRACT

Currently, manual posterior-anterior spinal stiffness assessment is widely used in daily practice to help determine how interventions for low back pain (LBP) are delivered. Given the poor reliability and validity of this assessment technique, mechanical indentation devices were developed. Early studies with these devices have shown an association between objective measures of stiffness and patient-reported outcomes. Unfortunately, these devices are typically designed to quantify stiffness only at one site which may ignore important data from other spinal segments. While it is possible to use these devices to obtain individual measurements from multiple sites, this approach requires time-consuming repositioning. To address this issue, a novel device called the VerteTrack (VT) was developed.

Use of spinal stiffness derived from VT may improve the generally poor performance of predictive models which are used in attempts to optimize treatment outcomes for LBP. As instrumented measurements of spinal stiffness have not been incorporated into these prediction models to date, their use may improve specific patient outcomes such as LBP disability.

This doctoral dissertation introduced and standardized a new spinal stiffness measurement device (VT) to provide data for predicting how patients respond to various interventions for LBP. Four studies informed the planning of this dissertation: 1) a reliability study to determine the within- and between-session reliability of lumbar stiffness measurements in asymptomatic participants using VT; 2) a Delphi study to develop a standard protocol for evaluating spinal stiffness and to improve the consistency of this assessment in future studies using the VT; 3) a secondary analysis

of a large RCT to determine if prediction of short-term treatment response can be improved by including spinal stiffness measures (i.e. lumbar spine stiffness) and 4) another secondary analysis of the same RCT to determine if prediction of long-term treatment response can be improved by including spinal stiffness measures (i.e. lumbar spine stiffness).

For Objective 1, the within and between-session reliability of lumbar spine stiffness measures using VT at the maximal tolerable load was excellent ranging from 0.95–1.00 and good to excellent ranging from 0.82–0.93, respectively.

Using a standard Delphi methodology, Objective 2 developed a consensus-based protocol for measuring spinal stiffness. In total, the pre-defined consensus threshold was reached for 67.2% (123/ 183) of statements after three rounds of surveys.

For Objective 3, a predictive model was developed for treatment response (30% improvement in Oswestry Disability Index) in a large RCT after 1 week. Response to treatment was predicted by a model containing height, gender, neck or upper back pain, pain frequency in the past 6 months, STarT Back Tool scores, patients' expectations about medication and strengthening exercises, and extension status. The model performed superiorly compared to prior predictive models, but spinal stiffness was not included in the final model.

A second predictive model was then created for Objective 4 using data from the same RCT. In this analysis, a novel approach was used that considered when responder status was first achieved during the trial (Response Onset (RO)), as well as if responder status was sustained (Response

Persistence (RP)). Baseline variables that univariately differentiated category membership in RO and RP groupings included pain frequency, depression, neck/upper back pain history, pain intensity, weight, spinal stiffness with STarT Back scores being specific to RO and ultrasonic muscle thickness measurements being specific to RP. Regression analysis predicted category membership correctly 46.1% for RO and 39.4% correctly for RP. Maximum terminal stiffness, pain frequency, and neck/upper back pain history appeared in both regression models with lumbar flexion and predicted success with stabilization exercises appearing only in the RO model.

To conclude, the VT is a reliable assessment device capable of measuring spinal stiffness continuously over an entire spinal region. A consensus-based protocol for measuring spinal stiffness using the VT is now available for operators to follow. Spinal stiffness measurements as collected by VT were not important in predicting treatment response in the short term but were a factor in identifying responders when multiple time point measurements were considered.

In total, this dissertation suggests spinal stiffness measured by VT has predictive values when long-term and multiple time-points assessments are considered. Future studies would ideally evaluate responder status at different time points to develop a full understanding of the stiffness phenomenon.

PREFACE

This dissertation is an original work completed by Maliheh Hadizadeh under the supervision of Greg N. Kawchuk, Professor in the Department of Physical Therapy at the University of Alberta. The first research project in this dissertation received research ethics approval from the University of Alberta Health Research Ethics Board, Pro00061205, 18 January 2016. The second research project in this dissertation received research ethics approval from the University of Alberta Health Research Ethics Board, Pro00102734, 20 August 2020. For the first and second projects, I was responsible for the design, data acquisition, data analysis, and preparation of the final reports. The third and fourth projects were secondary analyses of a large research collaboration, led by Dr. Greg Kawchuk at the University of Alberta and Dr. Julie Fritz at the University of Utah. Ethics approval was received from the University of Utah (IRB_00092127, 28 July 2016) and University of Alberta (Pro00067152, 15 September 2016), Institutional Review Boards. For the third and fourth projects, I was responsible for the data acquisition at Alberta site, data analysis, and preparation of the final reports.

Findings from the current doctoral dissertation have been published in peer-review journals:

- Chapter 3 of this dissertation has been published as: Hadizadeh M., Kawchuk G., Parent E. "Reliability of a new loaded rolling wheel system for measuring spinal stiffness in asymptomatic participants", BMC Musculoskeletal Disorders Journal. 2019; 20(1):176.
- Chapter 4 of this dissertation has been published as: Hadizadeh M., Kawchuk G., French S. "A consensus approach toward the standardization of spinal stiffness measurement using a loaded rolling wheel device: results of a Delphi study", BMC Musculoskeletal Disorders journal, 2021; 22(1):436.

- Chapter 5 of this dissertation has been published as: Hadizadeh M., Kawchuk G., Prasad N., Fritz J. “Predicting who responds to spinal manipulative therapy using a short-time frame methodology: results from a 238-participant study.”, PLOS One Journal, 2020, 15 (11): 1-23.

DEDICATION

I dedicate this work:

To my parents, who their love, patience, and support gives me the courage to challenge myself and allows me to realize my own potential and to pursue my goals.

To my beloved grandpa and grandma, to whom I could not say goodbye in person. Rest in Peace, I love and miss you so much.

To my partner, *Linden,* for his kind care, understanding, and for giving me motivation especially in the final stages of this thesis. I am truly thankful for having you in my life.

_Maliheh

ACKNOWLEDGMENTS

I would like to first say a big thank you to my supervisor Dr. Greg Kawchuk for all the support, feedback, and guidance he gave me throughout my Ph.D. program. I am very grateful to have been introduced to and get to work with high-level international collaborators through him and his projects. I further want to thank him for giving me the opportunity to attend several national and international conferences and share my work with scientists all around the world. Without his guidance and constant feedback, this Ph.D. would not have been achievable.

I would like to express my sincere gratitude to Dr. Narasimha Prasad for his guidance and kind advice in the statistical techniques applied in chapters 5 and 6 of this dissertation. I am also very grateful to my thesis committee members, Dr. Eric Parent and Dr. Jason Carey for their valuable contributions to this work. I am truly fortunate to have been academically supported by these talented and brilliant faculty members at the University of Alberta who gave me scientific guidance, many insightful suggestions and demonstrated a sincere interest in my work throughout this doctoral program.

Many thanks to Dr. Randy Vollrath, Dr. Shannon Wandler, Dr. Moe Gebara, Dr. Lacina Barsalou, Lynne Wong, and Erica Marr for their assistance during the data collection phase for chapter 5 and 6 of this dissertation. Working with these wonderful people made this more than a professional experience and motivated me to go on with the long process of data collection. Many thanks for all the nice discussions and moments shared together. A very special thanks to Dr. Julie Fritz, College of Health, University of Utah for giving me the opportunity to be a part of her big project and for her constructive comments which helped me to do better.

I gratefully acknowledge the Faculty of Rehabilitation Medicine for the financial support I was so fortunate to receive. I was supported by four-years Research Assistantship, Physical Therapy Graduate Student Fellowship Award, and twice nomination to the Faculty of Graduate Studies and Research for the Strathcona Physiotherapy Research Award. My gratitude is also owed to the Faculty of Graduate Studies and Research for supporting me financially through the Doctoral Recruitment Scholarship, Strathcona Physiotherapy Research Award, and the Graduate Travel Award. My gratitude also goes to Graduate's Students Associations for two Academic Travel Awards and the National Institute of Health for funding the final project that I was a part of.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

ACTex	Activating Exercise
BQ	Bournemouth Questionnaire
CPR	Clinical Prediction Rule
CI	Confidence Interval
FABQ	Fear-Avoidance Beliefs Questionnaire
FD curve	Force-Displacement Curve
HVLA	High-Velocity, Low-Amplitude
ICC	Intraclass Correlation Coefficient
LBP	Low Back Pain
LM	Lumbar Multifidus
MOBex	Mobilizing Exercise
MRI	Magnetic resonance imaging
N/A	Not Applicable
N	Newton
NPRS	Numerical Pain Rating Scale
N/mm	Newton/millimetre
ODI	Oswestry Disability Index
OR	Odds Ratio
PA	Posteroanterior
PGIC	Patient Global Impression of Change
PROMs	Patient-Reported Outcome Measures
RCT	Randomized controlled trial
REDCap	Research Electronic Data Capture
RMDQ	Roland Morris Disability Questionnaire
RO	Response Onset
ROC	Receiver Operating Characteristic
ROM	Range of Motion
RP	Response Persistence
SBT	The STarT Back Tool
SD	Standard Deviation
SEM	Standard Error of Measurement
SLR	Straight Leg Raise
SMT	Spinal Manipulative Therapy
SPSS	Statistical Package for the Social Science
UWCAP	University of Washington concerns about pain
UWPRSE	University of Washington pain-related self-efficacy
VT	VerteTrack

Chapter 1. General Overview of the Dissertation

1.1 Introduction

Low back pain (LBP) is usually defined by the location of pain or discomfort, typically below the costal margin and above the inferior gluteal folds, with or without referred leg pain (Hartvigsen et al., 2018). It is not a specific disease, rather it is a symptom that may result from a variety of different known or unknown abnormalities or diseases (Hartvigsen et al., 2018). Low back pain is classified as acute when it persists less than 6 weeks, subacute between 6 and 12 weeks, and chronic when it lasts longer than 12 weeks (Koes et al., 2006; Van Tulder et al., 2006). Most LBP is short-term and tends to recover on its own with no residual loss of function. Approximately one in five patients, however, develop chronic pain and disability for one year or more (Fourney et al., 2011).

For up to 90% of people presenting with LBP, the specific cause of the pain cannot be clearly identified and those are then classified as having so-called non-specific LBP (Maher et al., 2017). There are some underlying serious pathologies such as tumor or metastasis, visceral disease, vertebral compression fracture, infection, or inflammatory disorders that require diagnostic investigation and specific management of the condition, but these account for a very small proportion of LBP cases (Hartvigsen et al., 2018; Maher et al., 2017). Although the majority of LBP is thought to be mechanical in nature, most investigators would agree that mechanical back pain does not arise from a single cause but is a constellation of heterogeneous etiologies given the diversity of spinal tissues involved in the mechanical function of the spine (Fourney et al., 2011).

By extension, it is unlikely that these heterogeneous causes would respond uniformly and optimally to a specific treatment intervention.

Currently, clinical palpation technique often referred to as manual posterior-anterior (PA) spinal stiffness assessment is widely used in daily practice to target the treatment to the appropriate spinal segment and determine how interventions for LBP are delivered. The PA technique involves applying a manual PA force to the lumbar spine, typically spinous processes with the patient in a prone or seated position (Maitland et al., 2005). The clinician then subjectively interprets the resultant resistance to displacement and identifies segments of the spine with abnormal mobility and pain (Latimer et al., 1996; Maitland et al., 2005). Given the poor reliability ($ICC_{2,1}$: <0.4) and high variability of the practitioner-judged stiffness assessment technique (Maher & Adams, 1994), mechanical indentation devices were developed as an objective alternative. Such devices typically record the PA force-displacement data using a load cell attached to a computer-controlled stepper motor that indent the targeted site of spine or adjacent tissues with human operation (Owens et al., 2007). Development of these mechanical devices over the past 20 years helped researchers to obtain quantified measures of spinal stiffness. Early studies with these devices have shown an association between objective measures of stiffness and patient reported outcomes (Brodeur & Delre, 1999; Latimer et al., 1996). Unfortunately, these devices are typically designed to quantify stiffness only at one site which may ignore important data from other spinal segments. While it is possible to use these devices to obtain individual measurements from multiple sites, this approach requires time consuming repositioning. To address this issue, a novel device called the VerteTrack (VT) was developed.

1.2 The VerteTrack: A new device for spinal stiffness assessment

The VerteTrack device was developed in 2015 by Dr. Greg Kawchuk and his research team in the Rehab Robotics lab at the University of Alberta. We aimed to improve our prior spinal indentation (Wong et al., 2013) through continuous measure of the PA bulk deformation of the spine using a loaded wheel system. The VerteTrack not only offer stiffness data at one segment at a time but the whole spine in a fully computerized strategy rapidly and without the need for repositioning and recalibration between measurements. This continuous stiffness testing employs a loaded wheel system that moves uninterrupted over the spine while measuring the resulting load-displacement values along a subject-specific, laser-defined trajectory (Figure 1-1).

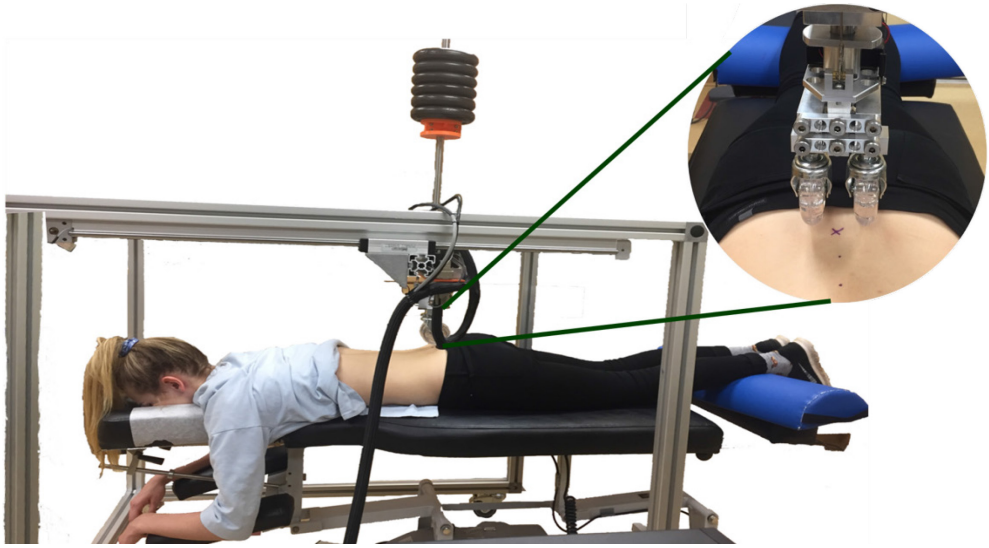


Figure 1-1 Spinal stiffness measurement by The VerteTrack device.

1.3 The clinical importance of spinal stiffness assessment

The PA spinal stiffness measurement has shown some promise as either a diagnostic tool or an outcome measure in LBP studies. In an animal model of lumbar spine degeneration, researchers found that spinal stiffness measurements are sensitive to disk lesions (Kawchuk et al., 2001). In a

cadaver model, PA spinal stiffness correlated with direct measurement of intersegmental flexibility in flexion/extension movements in the midthoracic spine (Sran et al., 2005). In a longitudinal study, spinal stiffness was found to decrease by 8% (1.2 N/mm) when patients were no longer in pain whereas asymptomatic control group showed no changes in PA stiffness over time (Latimer et al., 1996).

The manual assessment of spinal stiffness provides a basis for clinical decision-making process. Specifically, if a patient is judged with high lumbar spine stiffness (hypomobility) and no contraindications to mobilization/manipulation are present, mobilization or manipulation techniques are suggested (Fritz, Whitman, et al., 2005; Grieve, 1989; Maitland et al., 2005). On the other hand, a stabilization exercise program may be recommended when clinicians find hypermobility or low stiffness in the lumbar region (Fritz, Whitman, et al., 2005; Grieve, 1982; Paris, 1985). The PA spinal stiffness assessment has also shown some predictive validity in determining who is likely to respond best to different treatments (Koppenhaver et al., 2014). A finding of PA spinal stiffness assessment is 1 of the 5 variables formed a clinical prediction rule that was predictive of a successful reduction in disability with Spinal Manipulative Therapy (SMT) (Flynn et al., 2002). A randomized controlled trial validated this clinical prediction rule and its usefulness in predicting which patients with LBP are most likely to benefit from SMT (Childs et al., 2004). Spinal stiffness was also identified as a predictive factor of radiographic lumbar segmental instability (Fritz, Piva, et al., 2005) and reduction in disability with a standardized stabilization exercise program (Hicks et al., 2005). Furthermore, self-reported spinal stiffness was found as a potential key descriptor independent of pain in a sample of community-dwelling, older adults with LBP that may help to explain physical health and LBP-related disability in this

population (Sions & Hicks, 2017). This finding suggests that clinicians who are only treating pain may be missing an opportunity to improve clinical outcomes and enhance the efficiency of care by addressing spinal stiffness concurrently (Sions & Hicks, 2017).

1.4 Prediction of patients' outcomes

With the establishment of evidence-based treatments and developing health interventions towards personalized medicine strategies for many conditions, there is raising interest and need for clinical research on prediction of outcome for LBP at the early stage of treatment or even before the treatment begins (Mendonça et al., 2018). Prediction models are designed to help both clinicians and patients in making informed decisions about the use of diagnostic testing, starting/stopping/extending treatments, or making lifestyle changes (Harrell, 2015). These models can provide imperative insight into aspects that affect outcomes and costs of treatments and therefore can be beneficial for health care system as well as society (Bremer et al., 2018). Use of spinal stiffness derived from VT may improve the generally poor performance of predictive models which are used in attempts to optimize treatment outcomes for LBP. As instrumented measurements of spinal stiffness have not been incorporated into these prediction models to date, their use may improve specific patient outcomes such as LBP disability.

1.5 Primary aim and specific dissertation objectives

Given the above, the overall objective of this doctoral dissertation was to introduce and standardize a new spinal stiffness measurement device (VT) to provide data for predicting how patients respond to various interventions for LBP. Particularly, this dissertation had four specific objectives: 1) to determine the within- and between-session reliability of lumbar stiffness

measurements in asymptomatic participants using VT; 2) to develop a standard protocol for evaluating spinal stiffness and to improve the consistency of this assessment in future studies using the VT; 3) to determine if prediction of short-term treatment response can be improved by including spinal stiffness measures (i.e. lumbar spine stiffness) and 4) to determine if prediction of long-term treatment response can be improved by including spinal stiffness measures (i.e. lumbar spine stiffness).

1.6 Dissertation format

To achieve these objectives, two experiments (i.e., a reliability study and a Delphi study) and two secondary analyses were conducted (Figure 1-2). These studies are detailed in chapters 3 to 6 of this dissertation. The second chapter of this dissertation presents a review of what has been described in the scientific literature regarding relevant topics involved in this dissertation: Epidemiology of back pain and disease burden, lumbar posteroanterior segmental stiffness assessment, and prediction of intervention outcome for LBP.

Study 1 Chapter 3	To determine the within- and between-session reliability of lumbar stiffness measurements in asymptomatic participants using the new loaded rolling wheel system (VT).	BMC
Study 2 Chapter 4	To develop an updated best-practice protocol for evaluating spinal stiffness in human participants using the VT to improve the consistency of this assessment in future studies.	BMC
Study 3 Chapter 5	To determine if the baseline prediction of SMT responders can be improved through the use of stiffness data collected by VT.	PLOS
Study 4 Chapter 6	To determine the different recovery trajectories in a large RCT then determine if using participants' baseline characteristics including spinal stiffness measured by VT can predict the resulting trajectories.	

Figure 1-2 A summary of studies' objectives included in the current dissertation.

Chapter three describes findings from the first experiment that was published by Hadizadeh M., Kawchuk G., Parent E. "Reliability of a new loaded rolling wheel system for measuring spinal stiffness in asymptomatic participants", BMC Musculoskeletal Disorders Journal. 2019; 20(1):176. The protocol of this study was presented on the research day at the Faculty of Rehabilitation Medicine at the University of Alberta on June 3, 2016. The results were also presented at DC 2017 conference in Washington (15-18 March 2017), research day at the Faculty of Rehabilitation Medicine, the University of Alberta (9 June 2017), and the international society for the study of the lumbar spine, 45th ISSLS annual meeting in Banff (14-18 May 2018). This chapter presents the within- and between-session reliability of spinal stiffness measurements using VT, loaded versus unloaded conditions, and changes in measurement error by multiple trials. Furthermore, it introduces some parameters that should be included in future studies examining spinal stiffness in patients with spinal disorders.

Chapter four describes findings from the second study that was recently published by Hadizadeh M., Kawchuk G., French S. “A consensus approach toward the standardization of spinal stiffness measurement using a loaded rolling wheel device: results of a Delphi study”, BMC Musculoskeletal Disorders journal, 2021; 22(1):436. This Delphi study was built on current protocols and operators' expertise to develop an updated best-practice protocol for evaluating spinal stiffness in human participants using VT. It was hypothesized that expert feedback based on knowledge gained through training, experiences, and the VT operations manual would provide an agreement about standardizing the protocol of spinal stiffness measurements using the VerteTrack device. Based on the results of this study, a standardized protocol was established and available for researchers to evaluate spinal stiffness in future studies.

Chapters five and six describe the results from two secondary analyses of a large low back pain RCT. Chapter five aimed to determine if the baseline prediction of short-term treatment response can be improved by including spinal stiffness data collected by the VT device. Given the changes in spinal stiffness values following therapy, it was expected that this study would provide new insight into the predictive values of these variables. The results of this study were published by Hadizadeh M., Kawchuk G., Prasad N., Fritz J. “Predicting who responds to spinal manipulative therapy using a short-time frame methodology: results from a 238-participant study.”, PLOS One Journal, 2020, 15 (11): 1-23. The findings were also presented in CARLOQUIUM 2021 (virtual Conference) in 2-3 March 2021. This work was accepted for to be presented in the Spine Week 2020 meeting, in Melbourne on 27 April-1 May 2020 but the conference was cancelled because of pandemic. The study described in chapter 6 aimed to determine when response to treatment was achieved in a large RCT and then determined if using participants’ baseline characteristics can

predict the resulting responses. The results of this study were presented in WFC2021 Biennial congress in September 23-25. Finally, a synthesis of the above studies and directions for future research are presented in chapter seven of this dissertation.

1.7 References

Bremer, V., Becker, D., Kolovos, S., Funk, B., Van Breda, W., Hoogendoorn, M. & Riper, H. (2018). Predicting therapy success and costs for personalized treatment recommendations using baseline characteristics: Data-driven analysis. *Journal of Medical Internet Research*, *20*(8), 1–11. https://doi.org/10.2196/10275

Brodeur, R. R. & Delre, L. (1999). Stiffness of the Thoracolumbar Spine for Subjects with and without Low Back Pain. *JNMS - Journal of the Neuromusculoskeletal System*, *7*(4), 127–133.

Childs, J. D., Fritz, J. M., Flynn, T. W., Irrgang, J. J., Johnson, K. K., Majkowski, G. R. & Delitto, A. (2004). A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: A validation study. *Annals of Internal Medicine*, *141*(12), 920–928. https://doi.org/10.7326/0003-4819-141-12-200412210-00008

Flynn, T., Fritz, J., Whitman, J., Wainner, R., Magel, J., Rendeiro, D., Butler, B., Garber, M. & Allison, S. (2002). A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine*, *27*(24), 2835–2843. https://doi.org/10.1097/00007632-200212150-00021

Fourney, D. R., Andersson, G., Arnold, P. M., Dettori, J., Cahana, A., Fehlings, M. G., Norvell, D., Samartzis, D. & Chapman, J. R. (2011). Chronic Low Back Pain. *Spine*, *36*(21), S1–S9. https://doi.org/10.1097/BRS.0b013e31822f0a0d

Fritz, J. M., Piva, S. R. & Childs, J. D. (2005). Accuracy of the clinical examination to predict radiographic instability of the lumbar spine. *European Spine Journal*, *14*(8), 743–750. https://doi.org/10.1007/s00586-004-0803-4

Fritz, J. M., Whitman, J. M. & Childs, J. D. (2005). Lumbar spine segmental mobility assessment: An examination of validity for determining intervention strategies in patients with low back

pain. *Archives of Physical Medicine and Rehabilitation*, *86*(9), 1745–1752. https://doi.org/10.1016/j.apmr.2005.03.028

Grieve, G. (1982). Lumbar instability. *Physiotherapy*, *68*, 2–9.

Grieve, G. (1989). *Common vertebral joint problems*. Edinburgh: Churchill Livingstone.

Harrell, F. E. (2015). Regression modeling strategies with applications to linear models, logistic regression and survival analysis. In *Springer* (2nd ed.). https://doi.org/10.1198/tech.2003.s158

Hartvigsen, J., Hancock, M. J., Kongsted, A., Louw, Q., Ferreira, M. L., Genevay, S., Hoy, D. & Karppinen, J. (2018). *Series Low back pain 1 What low back pain is and why we need to pay attention*. *391*. https://doi.org/10.1016/S0140-6736(18)30480-X

Hicks, G. E., Fritz, J. M., Delitto, A. & McGill, S. M. (2005). Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization exercise program. *Archives of Physical Medicine and Rehabilitation*, *86*(9), 1753–1762. https://doi.org/10.1016/j.apmr.2005.03.033

Kawchuk, G. N., Kaigle, A. M., Holm, S. H., Fauvel, O. R., Ekström, L. & Hansson, T. (2001). The diagnostic performance of vertebral displacement measurements derived from ultrasonic indentation in an in vivo model of degenerative disc disease. *Spine*, *26*(12), 1348–1355. https://doi.org/10.1097/00007632-200106150-00018

Koes, B. W., Van Tulder, M. W. & Thomas, S. (2006). Diagnosis and Treatment of Low Back Pain. *Bmj*, *332*(7555), 1430–1434. https://doi.org/10.1136/bmj.332.7555.1430

Koppenhaver, S. L., Hebert, J. J., Kawchuk, G. N., Childs, J. D., Teyhen, D. S., Croy, T. & Fritz, J. M. (2014). Criterion validity of manual assessment of spinal stiffness. *Manual Therapy*, *19*(6), 589–594. https://doi.org/10.1016/j.math.2014.06.001

Latimer, J., Lee, M., Adams, R. & Moran, C. (1996). An Investigation of the Relationship between Low Back Pain and Lumbar Posteroanterior Stiffness. *Journal of Manipulative & Physiological Therapeutics*, *19*(9), 587–591.

Maher, C. & Adams, R. (1994). Reliability of Pain and Stiffness Assessments in Clinical Manual Lumbar Spine Examination. *Phys Ther*, *74*(9), 801–809. https://doi.org/10.1093/ptj/74.9.801

Maher, C., Underwood, M. & Buchbinder, R. (2017). Non-specific low back pain. *The Lancet*, *389*(10070), 736–747. https://doi.org/10.1016/S0140-6736(16)30970-9

Maitland, G., Hengeveld, E., Banks, K. & English, K. (2005). *Maitland's vertebral manipulation* (7th ed.). Philadelphia: Elsevier.

Mendonça, L., Monteiro-Soares, M. & Azevedo, L. F. (2018). Prediction of clinical outcomes in individuals with chronic low back pain: A protocol for a systematic review with meta-analysis 11 Medical and Health Sciences 1117 Public Health and Health Services. *Systematic Reviews*, *7*(1), 1–6. https://doi.org/10.1186/s13643-018-0818-2

Owens, E. F., DeVocht, J. W., Gudavalli, M. R., Wilder, D. G. & Meeker, W. C. (2007). Comparison of posteroanterior spinal stiffness measures to clinical and demographic findings at baseline in patients enrolled in a clinical study of spinal manipulation for low back pain. *Journal of Manipulative and Physiological Therapeutics*, *30*(7), 493–500. https://doi.org/10.1016/j.jmpt.2007.07.009

Paris, S. (1985). Physical signs of instability. *Spine*, *10*, 277–279.

Sions, J. M. & Hicks, G. E. (2017). Back Stiffness Is Associated with Physical Health and Low Back Pain-Related Disability in Community-Dwelling Older Adults. *Pain Medicine (Malden, Mass.)*, *18*(5), 866–870. https://doi.org/10.1093/pm/pnw107

Sran, M. M., Khan, K. M., Zhu, Q. & Oxland, T. R. (2005). Posteroanterior stiffness predicts

sagittal plane midthoracic range of motion and three-dimensional flexibility in cadaveric spine segments. *Clinical Biomechanics*, *20*(8), 806–812. https://doi.org/10.1016/j.clinbiomech.2005.04.009

Van Tulder, M., Becker, A., Bekkering, T., Breen, A., Del Real, M. T. G., Hutchinson, A., Koes, B., Laerum, E. & Malmivaara, A. (2006). Chapter 3: European guidelines for the management of acute nonspecific low back pain in primary care. *European Spine Journal*, *15*(SUPPL. 2), 169–191. https://doi.org/10.1007/s00586-006-1071-2

Wong, A. Y. L., Kawchuk, G., Parent, E. & Prasad, N. (2013). Within- and between-day reliability of spinal stiffness measurements obtained using a computer controlled mechanical indenter in individuals with and without low back pain. *Manual Therapy*, *18*(5), 395–402. https://doi.org/10.1016/j.math.2013.02.003

Chapter 2. Literature Review

The purpose of this chapter is to present some information on study variables through a review of literature on the existing research that helps to understand the individual studies presented in the next chapters. As such, this chapter is presented in three main sections: The first section provides a broad overview of the epidemiology of back pain and disease burden. The second section is a more focused presentation of the lumbar PA segmental stiffness assessment. Finally, some information about the prediction of intervention outcome for LBP is introduced in the last section.

2.1 Epidemiology of Back Pain and Disease Burden

2.1.1 Burden and prevalence of back pain

Low back pain (LBP) is a common health condition globally (Hoy et al., 2012; Vos et al., 2016), with over half a billion people had LBP in 2015 (Hurwitz et al., 2018). Findings of the most recent Global Burden of Disease study highlighted the increased global prevalence of LBP by 17.3% from 2005 to 2015 (Vos et al., 2016). Among 315 diseases and injuries, low back and neck pain were ranked the fourth leading cause of disability-adjusted life years globally after ischemic heart disease, cerebrovascular disease, and lower respiratory infection. In contrast, low back and neck pain were ranked as the 12^{th} in 1990 and the 8^{th} in 2005 (Kassebaum et al., 2016). Similarly, they were identified as the leading causes of years lived with disability in most countries and age groups (Hurwitz et al., 2018; Vos et al., 2016)

A high prevalence of LBP is reflected in Canada as well (Angarita-Fonseca et al., 2019; Bath et al., 2014; Gross et al., 2006; Moulin et al., 2002; Schopflocher, 2003; Schopflocher et al., 2011), impacting one in five Canadians (Angarita-Fonseca et al., 2019). The low back was reported as the most frequent locus of chronic pain among Canadian adults populations with more than one-third of those suffering from chronic pain experienced LBP (Schopflocher et al., 2011). A population-based analysis of the 2009-2010 Canadian community health surveys also revealed that 20.2% of Canadian adults have back problems lasting 6 months or more (Bath et al., 2014). Similarly, the lifetime prevalence of back pain in Alberta and Saskatchewan was 83.8%, with 12.3% of respondents reporting taking time off from work (Gross et al., 2006).

2.1.2 Economic burden

Given the high prevalence of LBP, it is not surprising that it is associated with enormous health care costs (Dagenais et al., 2008; Geurts et al., 2018; Katz, 2006; Kim et al., 2019; Lim et al., 2006; Maher et al., 2017), mainly in terms of lost workdays and reduced productivity (Dagenais et al., 2008; Katz, 2006). In addition, there are considerable implications for the use of health care resources, particularly for chronic conditions, which are a major cause of long-term disability, distress, and work loss (Kay et al., 2017; Lim et al., 2006; Maher et al., 2017). Back and spine disorders cost Canada's health system between $6 and $12 billion annually (*Bone and Joint Canada. Low Back Pain*, 2013). It was identified as the highest overall degree of resource use with arthritis or rheumatism, high blood pressure, and migraines among people less than 60 years of age in Canada (Rapoport et al., 2004).

2.2 Lumbar posteroanterior segmental stiffness assessment

Prior studies have demonstrated that the mechanical properties of the spine such as segmental motion/ stiffness change in individuals with LBP (Colloca & Keller, 2001; Ferreira et al., 2009; Latimer, Lee, et al., 1996; Thakral et al., 2014; Wong & Kawchuk, 2017). Biomechanically, stiffness of a structure is defined as the extent to which the structure resists deformation in response to an applied force (Baumgart & Cordey, 2001). However, the relationship between subjective feelings and objective measurements of spinal stiffness is complicated. A prospective clinical study investigated the relationship between biomechanical back stiffness and the reported feeling of stiffness (Stanton et al., 2017). The results showed a conscious perception of feeling stiff does not reflect the actual biomechanical back stiffness. In other words, what people describe as a feeling of stiffness is different from the biomechanical tissue state. For some people, a feeling of stiffness reflects a perceived resistance to movement, for others, it's a feeling of a lack of movement velocity or even some describe it as fear of movement (Stanton et al., 2017). Therefore, the researchers suggested that the feeling of stiffness may represent a protective perceptual inference that is created by the nervous system to reduce movement and prevent any risk of further injuries (Stanton et al., 2017). In contrast, the most recent study suggests that the measurement and perception of stiffness may be more related to instrumented spinal stiffness measures if other factors such as age and sex are controlled for (Harsted et al., 2021). The focus of this chapter, however, is on the measurement of spinal stiffness, not the perception of spinal stiffness.

Methods for spinal stiffness assessment include both manual assessment (practitioner palpation and judgment) and instrumented assessment (mechanical devices).

2.2.1 Manual assessment

Clinically, spinal posteroanterior (PA) stiffness is described as the perceived resistance through the intervertebral joint range of movement during the application of manual forces in a PA direction to the skin overlying the spinous processes of a prone patient (Figure 2-1) (Maitland et al., 2005). Differences in the amount, behaviour, and quality of PA stiffness between adjacent vertebral levels help clinicians to identify symptomatic joints that may require treatment, prescribe the most appropriate treatment, and finally to assess the patient's responses to the treatment (Kenna & Murtagh, 1997; Maitland et al., 2005). This technique is commonly performed as part of patients' evaluation by physical therapists, osteopaths, chiropractors, as well as medical practitioners in the management of LBP (Owens et al., 2007). Results of the manual assessment are typically reported as if the evaluated segment is hypermobile, hypomobile or within normal limits (Maitland et al., 2005).

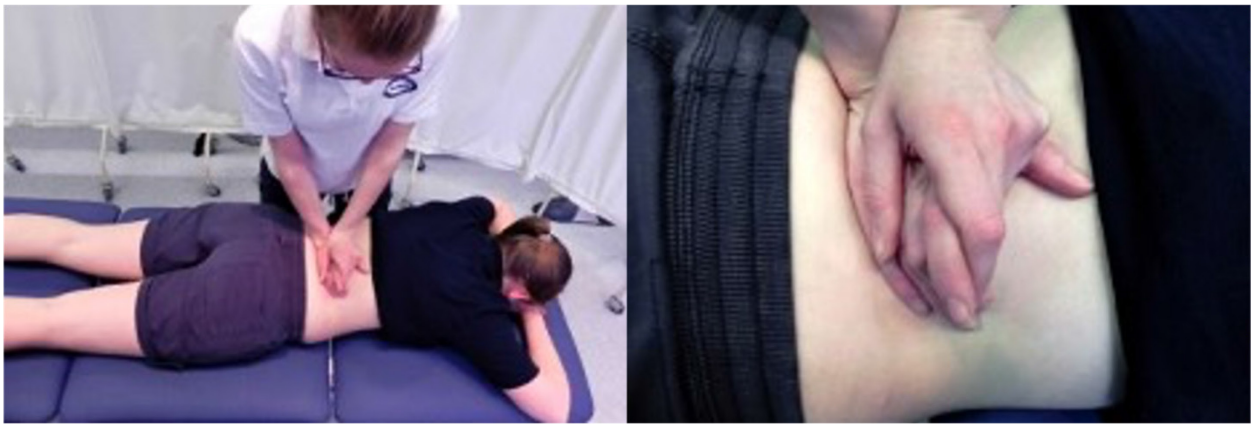


Figure 2-1 Manual posteroanterior segmental mobility assessment

2.2.2 Instrumented assessment

Various mechanical devices have been used for objective assessment of PA spinal stiffness. Dynamic PA spinal stiffness assessments were conducted in some studies using a handheld mechanical device (Activator Adjusting Instrument) equipped with an impedance head (load cell and accelerometer). Apparent mass measurements were then calculated (peak force/ peak acceleration, kg) as a measure of the dynamic spinal stiffness characteristics at the segmental contact points (Colloca et al., 2003; Colloca & Keller, 2001, 2004; Keller et al., 1999). However, most previous studies have reported the load deformation response by applying a standardized mechanical PA force to the selected vertebral level using a probe or indenter and quantifying the concurrent displacement (Figure 2-2 A) (Snodgrass et al., 2012). Some researchers proposed assisted devices which require the operator to use an indentation probe to manually apply the force to lumbar (Figure 2-2 B) (Owens Jr et al., 2007; Stanton & Kawchuk, 2009) and cervical spine segments (Tuttle et al., 2008a, 2008b).

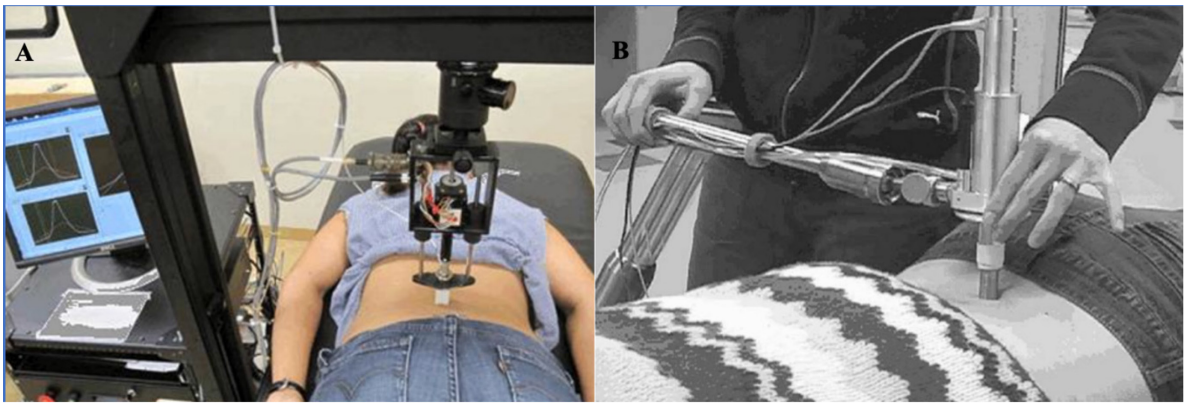


Figure 2-2 Lumbar Spine stiffness assessment by A. Mechanical indentation device and B. Assisted indenter

Following the collection of force and displacement data, spinal stiffness is calculated from the resulting Force-Displacement (FD) curve. The curve has two distinct parts: (1) the non-linear part or the toe region and (2) the linear part (Latimer, Goodsel, et al., 1996). The toe region represents the initial loading phase where small forces applied to the spine generate relatively large displacements while the linear part of the curve indicates the resistance perceived after the initial loading phase for forces over 20-30 Newtons (N) (Latimer, Goodsel, et al., 1996). Different features of the FD curve have been used in the literature to quantify PA segmental movements including the toe region length, the displacement magnitude at 30 N, the linear region's slope, the final loading force, and the overall displacement (Wong & Kawchuk, 2017). However, two of these features seem to be more practical: the slope of the FD curve between 5 N and 60 N known as global stiffness and the ratio between the maximal applied force to the maximal resultant displacement known as terminal stiffness, both are reported in N/mm (Figure 2-3) (Fritz et al., 2011; Stanton & Kawchuk, 2009; Vaillant et al., 2010; Wong et al., 2013).

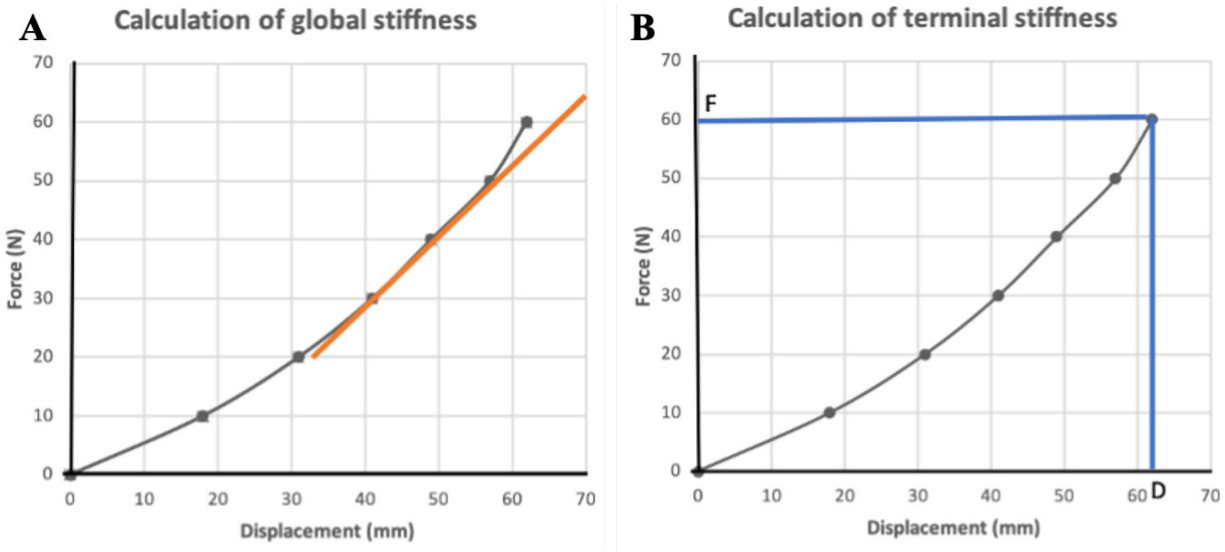


Figure 2-3 Force-Displacement curve. A. Calculation of global stiffness and B. Calculation of terminal stiffness.

2.2.3 Psychometric properties of stiffness assessments

Although there is some evidence in favor of the clinical validity of manual PA segmental stiffness assessment (Abbott et al., 2009; Brunarski, 1982; Fritz et al., 2005; Harvey & Byfieid, 1991; Jenson et al., 1993), a recent study (Koppenhaver et al., 2014) found no correlation between manual assessment of spinal stiffness and a criterion measure using spinal indentation in patients with LBP. Similarly, sensitivity and specificity estimates of judgments of hypomobility were low (0.20–0.45) and positive and negative likelihood ratios were not statistically significant (Koppenhaver et al., 2014).

Reliability of manual assessments of PA segmental stiffness has been largely studied. Systematic reviews reported poor to fair inter-examiner reliability and moderate to substantial intra-examiner reliability for practitioners' stiffness judgments made with this technique (Hestbœk & Leboeuf-Yde, 2000; Huijbregts, 2002; Seffinger et al., 2004; Snodgrass et al., 2012; Stochkendahl et al., 2006; Van Trijffel et al., 2005; Wong & Kawchuk, 2017). Factors affecting practitioner judgment of stiffness are categorized into three different domains: examiner-related factors (visual occlusion, hand contact area, hand position, and the magnitude/ frequency/speed/angle of force applied, identification of a specific spinous process as a PA pressure target), patient-related factors (gender, skinfold, fat composition, testing position, trunk muscle contractions, intra-abdominal pressure, respiratory cycle), and environmental factors (padding of test surface) (Snodgrass et al., 2012; Wong & Kawchuk, 2017). Most of these factors seem to be controllable and multiple recommendations have been suggested to enhance the reliability of this technique, however, the beneficial effects of these recommendations have not yet been examined (Wong & Kawchuk, 2017). The reliability of mechanically assisted devices indicated considerably better results than

that reported for purely manual assessments (intraclass correlation coefficient (ICC)> 0.75) (Owens Jr et al., 2007; Stanton & Kawchuk, 2009; Tuttle et al., 2008a).

Studies reported high test-retest reliability of instrumented spinal stiffness measurements. According to a recent narrative review on manual and instrumented methods, within- day ICCs of instrumented spinal stiffness measurements ranged from 0.79 to 0.99, whereas the between-day ICCs ranged between 0.88 and 0.98 (Wong & Kawchuk, 2017). The patient-related factors and environmental factors mentioned above can also affect the reliability of the instrumented spinal stiffness assessments.

2.2.4 Clinical application

The association between biomechanical back stiffness and patient-reported outcome measures is complex. While scientists have tried to measure stiffness using mechanical devices that quantified stiffness through measuring the load-displacement behavior of lumbar motion segments, the measure did not correlate with the reported feeling of stiffness. In another clinical study, researchers recruited one hundred and ninety-one patients with chronic LBP to explore the relationship between manual assessment of PA spinal stiffness and self-reported outcome measures including patient-specific functional status, global perceived effect, pain, and disability (Ferreira et al., 2009). The patients were randomly allocated into three treatment groups: spinal manipulative therapy, motor control exercise, or a general exercise program. Spinal stiffness was manually assessed before and after 8 weeks of treatment along with other clinical outcomes. The results showed a decrease in spinal stiffness following treatment, more so in those with the stiffest spines meaning that the change in stiffness is negatively correlated with initial stiffness. However,

no significant association was observed between initial PA stiffness and any of the final outcome measures (Ferreira et al., 2009). In addition, the association of lumbar spine stiffness with patient-reported pain and disability in adults with chronic LBP who received 12 sessions of spinal manipulative therapy (SMT) over 6 weeks were examined in a single-arm clinical trial (Xia et al., 2017). In this study, the global lumbar spine stiffness was obtained at L3 using three methods: hand palpation, a hand-held instrumented device, and an automated indenter device. The results revealed higher levels of hand-held and automated stiffness measures were significantly associated with higher levels of disability. However, no association was found between lumbar spine stiffness and pain intensity (Xia et al., 2017). In a monocentric, individually controlled, experimental trial, myofascial tissue stiffness was measured in 40 patients with chronic neck and back pain (Lederer et al., 2019). Researchers hypothesized the most painful region in the neck or lower back might be the segment with the highest stiffness. Therefore, the more painful side was treated with a cupping massage while the contralateral side was served as an individual control. Tissue stiffness was then measured using a small hand-held device called myometer before and after the treatment. Patients were asked to rate their pain on a standardized pain questionnaire before and 24 hours after treatment. Analysis of data indicated that the more painful side was not higher in stiffness compared to the contralateral control side before treatment. The tissue stiffness of the treated region decreased significantly after treatment but returned to baseline after 24 hours; at the same time, patients' pain ratings improved substantially.

Decrease in the mobility of the lumbar spine has been frequently reported among patients with LBP (Colloca & Keller, 2001; Ferreira et al., 2009; Latimer, Lee, et al., 1996; Thakral et al., 2014). Early research using a portable stiffness testing device reported an association between LBP and

PA spinal stiffness. Findings from this study showed that patients with LBP have a reduction in PA stiffness (14-37%) as their pain decreased (mean time between tests = 22.64 days, range = 2-105 days). This decrease was not observed in asymptomatic controls who were matched with the LBP group on gender, age, vertebral level tested and time between tests (Latimer, Lee, et al., 1996). A recent study (Stanton et al., 2017) revealed that a self-protective response existed in patients with chronic LBP who reported stiffness: These patients significantly overestimated force applied to their spine, yet were more sensitive at detecting changes in this force or feelings of back stiffness compared to healthy controls without LBP. This finding suggests that individuals with back pain may feel stiff as an effective perceptual mechanism to limit movement and thus avoid further injuries (Stanton et al., 2017).

In a prospective case series study, researchers evaluated associations between spinal stiffness characteristics measured by a mechanical device and clinical outcome in patients with nonspecific LBP following 2 SMT sessions (Fritz et al., 2011). They found a significant reduction in global stiffness immediately in response to SMT was associated with significant improvement on the Oswestry Disability Index (ODI) after one week. In another study with a similar methodology, significant immediate reduction in the average L3 segmental stiffness after each SMT session and sustained decreases in the average L3 segmental stiffness after 1 week was observed in patients with LBP who also showed clinically significant improvement in ODI. These studies suggest that an immediate post-SMT decrease in spinal stiffness might be an independent predictor of improved ODI scores at 1 week (Wong & Kawchuk, 2017).

A cohort study with 680 participants and 18-month follow up found that a self-reported feeling of stiffness is a significant predictor of mobility disability (Thakral et al., 2014). Due to this association, it is considered as an important indicator of function and quality of life especially for elderly adults and their healthcare providers (Thakral et al., 2014). Moreover, estimates of lumbar hypomobility made during manual physical assessment were identified as one of the five variables that predict which patient may respond favorably to SMT (Childs et al., 2004; Flynn et al., 2002)

In a randomized crossover study, researchers investigated the relation between spinal stiffness and LBP (Wong et al., 2016). In order to induce temporary pain, an equal volume of hypertonic or isotonic saline in random order was injected into the L3–L5 interspinous ligaments of nine asymptomatic participants in two separate sessions. Pain intensity, spinal stiffness at the L3 level, and the surface electromyographic activity of six trunk muscles were measured before, immediately after, and 25-minute after injections. Researchers observed temporary increases in spinal stiffness and concurrent trunk muscle co-contraction following the injections. This Finding support the role of spinal stiffness assessments in monitoring back pain progression (Wong et al., 2016).

2.2.5 Limitations of spinal stiffness assessment

Several challenges have been identified in measuring spinal stiffness. First, the stiffness judgments made with manual assessment highly depend on human performance and interpretation. As a result, this manual technique is limited in sense of human perception, poor reliability, and large variability. Second, although, the instrumented assessment of spinal stiffness has partially solved these issues and leads to more reliable measurements, it has presented a new set of challenges.

Mechanical devices that are used for spinal stiffness measurement are primarily designed for research applications. As a result, many features of these devices such as their size, cost, time consumption, and complex operation limit their implications in the clinical settings. Third, a numerous range of variables should be controlled in order to have reliable and valid measurements which is not always feasible to control all of these variables for routine use in practice. Forth, despite various mechanical devices being available in the research centers around the world for quantifying spinal stiffness, there is limited research on developing a standard operating protocol for these measurements. Only one study reported a standard protocol for a portable device_ therapeutic spinal mobilizer_ in a pilot sample of five participants. The authors recommended a protocol in which the operator requires using an optimum load of 90 N over the spinal segment of interest with a loading frequency between 0.5 and 0.1 Hz. The stiffness values are recorded in 3–5 cycles of loading with the participant in prone position on a standard plinth at the functional residual capacity during a respiratory cycle. They suggested following this protocol for measuring spinal stiffness using the therapeutic spinal mobilizer will yield stiffness values that can be compared between cases. Finally, the lack of a reference or gold standard measure of instrumented spinal stiffness assessment makes the comparison between the studies almost impossible. The accuracy of a diagnostic test is determined by comparing it to a gold standard which is a common and well-accepted method of identifying a disease or a clinical condition (Nordin et al., 2009).

Although a gold standard for spinal stiffness measurements is not well established, few studies investigated manual assessments of lumbar segmental mobility compared to lumbar segmental mobility assessed by radiographic measurements. Two previous studies found moderate agreement between manual assessments of intervertebral motion and segmental motion during flexion-

extension radiographs of the lumbar spine (Abbott et al., 2005; Fritz et al., 2005). Another study reported poor agreement in ratings of spinal motion between simultaneous manual and dynamic MRI assessments (Landel et al., 2008).

Overall, a clear understanding of the relevant characteristics of spinal stiffness not only has the potential to improve the effectiveness of manual therapy diagnosis and treatment, but to improve the efficiency of rehabilitation interventions in the management of LBP in both research and practice.

2.2.6 Previous studies on the VerteTrack device

The comfort and safety of the VT for the assessment of lumbar trunk stiffness in a sample of eighty-four young adults were investigated by a group of researchers at the Macquarie University (Brown et al., 2017). They reported minor and short-lived adverse events associated with the VT testing (2.4%) which involved two participants with pre-existing neuromusculoskeletal conditions. An analysis of the comfort data during and after the assessment indicated that the majority (75%) of participants in this study found the VT assessment comfortable. The comfort ratings were inversely related to loading; meaning that increasing loads (≥30 N) resulted in lower comfort ratings. Furthermore, those who experienced one or more days of LBP in the past week were more likely to report a lower comfort rating compared to asymptomatic individuals. Since tolerance for spinal stiffness testing appears to be individual in nature, researchers recommended that testing be performed to the participant's onset of discomfort rather than an absolute loading value (Brown et al., 2017).

Another experiment was conducted in 2018 at the Macquarie University to investigate the accuracy (precision and bias) of stiffness measurements obtained by the VT device (Young et al., 2020). Measurements were performed on a viscoelastic foam medium for both single-level and multiple-level continuous stiffness assessments. Given that no 'gold standard' exists to ascertain spinal stiffness in human participants and single-level indentation is the more established method of indentation reported in the literature, this method was selected to be used as a proxy reference standard in this study. The findings of this study revealed that the VT device has high accuracy (high precision, low systematic bias) under bench-top conditions compared to reference values (Young et al., 2020).

2.3 Prediction of intervention outcome for LBP

There is a number of information on predicting LBP patients' responses to various interventions before the application of any care. Different potential predictive factors have been measured in order to estimate the probabilities of outcomes. The focus of this chapter, however, is on studies in which potential predictors to SMT have been studied in patients with LBP at the baseline.

2.3.1 Clinical Prediction Models for SMT

Previous predictive studies examined a range of prognostic factors in relation to LBP patients' response to SMT in demographic, clinical history, patient-reported outcome measures (PROMs), and physical examination domains.

Individual characteristics (e.g., age, gender, race, and ethnicity) and social determinants of health (e.g., educational opportunities, income level, employment status, and having medical coverage) represent the demographic domain in clinical research. Individual characteristics in the demographic domain are commonly examined in studies predicting treatment outcomes in patients with LBP, however, social determinants of health are not often included. Evidence suggests that valuable insights can be gained from these predictors; Over a 25-day sick leave during the previous year was an independent predictor of poorer LBP treatment outcome in a prospective randomized trial (Niemistö et al., 2004), whereas being employed reduces the chances of poor outcomes on disability in a prospective cohort study (Cruz et al., 2020). A multicenter retrospective analysis study found that low education level negatively influenced the amount of perceived pain and disability in 310 outpatients with chronic non-specific LBP (Ferrari et al., 2019). A prospective study from the Nordic back pain subpopulation research program examined 50 potential baseline factors from demographic and clinical history in a sample of 875 LBP patients and reported a 5-variable model predicting chiropractic treatment outcome at the fourth visit including two demographic variables (women and some sort of social benefit) (Leboeuf-Yde et al., 2004).

Pain characteristics (e.g., pain severity, pain duration, anatomical location, and pain frequency) and health history that might be relevant to a patient's current health status (e.g., other prior/current pathologies, past treatments, medications the patient is taking or may have recently stopped taking, and past surgical history) represent the clinical history domain. The symptom duration has been reported in the number of studies as a baseline predictor for treatment outcome in LBP patients (Axén et al., 2005; Flynn et al., 2002; Langworthy & Breen, 2007; Leboeuf-Yde et al., 2004; Newell & Field, 2007; Skargren & Öberg, 1998). Other reported prognostic factors from this

domain include leg pain (Axén et al., 2005; Flynn et al., 2002; Malmqvist et al., 2008), the severity of symptoms (Leboeuf-Yde et al., 2004)(Leboeuf-yde & Larsen, 2005), and more than one site of pain (Skargren & Öberg, 1998) such as additional neck pain (Leboeuf-Yde et al., 2004). Adequate representation of the clinical history domain at baseline is essential to allow for clear determination of which characteristics of the pain experience and health history may have strong and consistent associations with treatment outcome in the short and long term (George et al., 2020).

Patient reported outcome measures are questionnaires patients complete to provide information on aspects of their health status such as symptoms, functionality, physical, mental, and social health and overall quality of life. Information from the patient's perspective can help to monitor individual patient progress, facilitate communication between professionals and patients, adjust treatment and care to ensure people are getting the most benefit from their care, and eventually to improve the quality of health care services (Canadian Institute for Health Information (CIHI)). Many different PROMs have been studied in the prediction of LBP outcomes, and they can be broadly categorized into the measure of disability, pain intensity, quality of life, expectations, and psychosocial factors.

Functional disability and pain intensity are consistently included in LBP prediction studies for measuring response to treatment. Measures of functional disability and pain intensity (a continuous measure) were used as outcomes in 8/23 (Burton et al., 1995, 2004; Cecchi et al., 2011; Childs et al., 2004; Flynn et al., 2002; Leboeuf-Yde et al., 2004; Leboeuf-yde & Larsen, 2005; Underwood et al., 2007) and 4/23 (Field & Newell, 2012; Leboeuf-Yde et al., 2004; Leboeuf-yde & Larsen, 2005; Vavrek et al., 2015) articles, respectively. Health-related quality of life measure is also well

established in the study of LBP, however, only one study examined this parameter in predicting response to SMT and they did not find any significant results (Underwood et al., 2007).

Patient's expectation is another potentially modifiable prognostic factor that has shown promising results in the LBP prognostic factor reviews (Fadyl & McPherson, 2008; Hayden et al., 2019; Iles et al., 2009), however, little focus has been given to this factor in studies predicting SMT outcomes. In particular, three types of expectations relevant to the LBP field are: General/recovery expectations, self-efficacy expectations, and treatment expectations. General expectations are defined as expectations that a clinical outcome will occur; self-efficacy expectations are a person's perceptions about their ability to perform behaviours to achieve a future outcome, and treatment expectations describe expectations about clinical outcomes specifically related to an ongoing treatment (Hayden et al., 2019). A recent study indicated that patients seeking care from chiropractors with a high expectation for recovery had a 58% greater chance to report a short-term improvement (Eklund et al., 2019). Researchers recommended that clinicians consider their patient's expectations at an early stage of treatment to identify those who are at risk of a poor prognosis (Eklund et al., 2019). Secondary analysis of the UK BEAM dataset suggested that in those allocated to combined treatment (manipulation and exercise), expecting treatment to be helpful might improve outcome at 12 months (Underwood et al., 2007). In addition, data from a randomized trial of patients with back/neck pain receiving either chiropractic or physiotherapy as primary management showed patients' expectations of treatment was one of five prognostic factors being significantly associated with Oswestry score at the 12-month follow-up (Skargren & Öberg, 1998). Beyond these studies, the PROMs domain has been largely unexplored in terms of patient's

expectations in studies predicting SMT outcomes. Including this in future studies might improve prediction accuracy.

The psychological aspect of PROMs has been frequently used in previous predictive studies to determine the overall level of distress associated with LBP including negative effects (e.g. depressive symptoms (Burton et al., 1995, 2004; Leboeuf-Yde et al., 2009; Newell & Field, 2007) and anxiety (Langworthy & Breen, 2007; Newell & Field, 2007)) and coping styles (e.g. fear-avoidance (Burton et al., 1995, 2004; Field et al., 2010; Flynn et al., 2002; Langworthy & Breen, 2007; Leboeuf-Yde et al., 2009; Newell & Field, 2007), pain catastrophizing (Burton et al., 1995; Field et al., 2010), and self-efficacy (Burton et al., 1995; Field et al., 2010)). However, the predictive value of psychological and behavioural variables in LBP patients receiving care from chiropractors appeared less important to the outcome than other factors. While an early prospective study showed psychosocial parameters could predict the 1-year disability in an LBP population seeking primary care from osteopaths (Burton et al., 1995), more recent studies have found little or no correlation with outcomes (Eklund et al., 2016; Field et al., 2010; Field & Newell, 2012; Langworthy & Breen, 2007; Leboeuf-Yde et al., 2009; Newell & Field, 2007). Although, various screening tools are available for measuring different features in this domain and there's no superior single measure to recommend, it seems important to ensure the measures used in the model are fully representative of the domain including both negative (e.g., fear-avoidance and catastrophizing) and positive (e.g. self-efficacy and acceptance) features (George et al., 2020).

The final domain to consider in LBP prediction studies is the physical examination domain. Specific measures recommended for this domain have included lumbar ROM (e.g. flexion,

extension, and left/right lateral bending), left/right hip external/internal rotation, provocation tests (e.g. posterior shear test), motion tests (e.g. Gillet test and seated flexion test), symmetry tests (e.g. posterior superior iliac spine symmetry in sitting and standing), special tests (e.g. centralization/peripheralization, left/right straight leg raise, and sit-up test), manual motion palpation (e.g. L1-L5 hypo or hypermobility), pain on palpation, pain pressure threshold. Variables obtained from physical examination seem important in predicting response to SMT as demonstrated in the clinical prediction rule for classifying LBP patients following SMT. The final model consists of 5 variables representing two from the physical examination including one hip with more than 35° of internal rotation range of motion and hypomobility in the lumbar spine (Childs et al., 2004; Flynn et al., 2002).

2.3.2 Limitations of predicting response to SMT

There were several limitations of included studies about predicting response to SMT in patients with back pain in this chapter. One of the limitations was the heterogeneity at various levels including study samples, population characteristics, interventions, outcome measures, and analytical approaches. For instance, the sample size varied from 69 (Cecchi et al., 2011) to 1116 (Underwood et al., 2007), studies' population was heterogenous meaning that studies included patients with LBP of any duration (chronic, subacute, acute, and persistent LBP) with or without leg pain and any intensity (mild, moderate, and severe), the duration of the interventions ranged from 1 week (Flynn et al., 2002) to 12 weeks (Underwood et al., 2007) with most studies taking a pragmatic approach and not reporting the exact duration of SMT. Studies had different periods of response assessment time from 1 week (Peterson et al., 2012) to 4 years (Burton et al., 2004). Most of the included studies predicted outcome post-treatment, but some also predicted outcome at 4-

weeks (Axén et al., 2005; Langworthy & Breen, 2007; Latimer, Lee, et al., 1996; Leboeuf-Yde et al., 2004; Skargren & Öberg, 1998), 6-weeks (Langworthy & Breen, 2007; Vavrek et al., 2015), 3-months (Cecchi et al., 2011; Field & Newell, 2012; Leboeuf-Yde et al., 2005, 2009; Newell & Field, 2007; Peterson et al., 2012; Underwood et al., 2007), 6-months (Cecchi et al., 2011; Childs et al., 2004; Leboeuf-Yde et al., 2005), 12-months (Burton et al., 1995; Cecchi et al., 2011; Eklund et al., 2019; Leboeuf-Yde et al., 2004, 2005; Leboeuf-yde & Larsen, 2005; Skargren & Öberg, 1998; Underwood et al., 2007; Vavrek et al., 2015), 18-months (Leboeuf-Yde et al., 2005) and 48-months (Burton et al., 2004) follow-ups. Some studies had multiple response assessment time and the variables that appeared in their model changed at different time points (Leboeuf-Yde et al., 2004; Peterson et al., 2012). Studies also used a wide variety of outcome measures. Eight studies used self-reported LBP status on a five-point Likert scale to measure patients' improvements (Axen et al., 2005; Eklund et al., 2016, 2019; Iben et al., 2002; Leboeuf-Yde et al., 2005, 2009; Malmqvist et al., 2008), four employed patient global impression of change (Field et al., 2010; Field & Newell, 2012; Newell & Field, 2007; Peterson et al., 2012), one study chose Deyo's core set (Langworthy & Breen, 2007), and the rest used self-reported pain and function measured by different tools such as the numeric pain rating scale (Leboeuf-Yde et al., 2004) (Leboeuf-yde & Larsen, 2005), modified Von Korff pain scale (Vavrek et al., 2015), Bournemouth questionnaire (Field & Newell, 2012; Newell & Field, 2007), Roland Morris disability questionnaire (Burton et al., 1995, 2004; Cecchi et al., 2011; Leboeuf-yde & Larsen, 2005; Underwood et al., 2007), and ODI (Childs et al., 2004; Flynn et al., 2002; Leboeuf-yde & Larsen, 2005; Skargren & Öberg, 1998). Although the choice of outcome measures depends on specific research goals, such variability prevents progress in identifying predictive factors that generalize to other outcomes relevant to the LBP patient's response to SMT. For example, a model designed specifically for

prediction of response to SMT based on pain intensity may not be well suited for predicting SMT responders based on disability or patient satisfaction.

Only one studies focused on only SMT (Flynn et al., 2002), two studies SMT and exercises (Childs et al., 2004; Underwood et al., 2007), and others reported chiropractic management as decided by the treating chiropractor as for the intervention which was a combination of three or more of the following interventions: general advice, positive encouragement, booklet, passive soft tissue stretching, passive articulation of the lumbar spine, direct and indirect mobilization, SMT with associated soft tissue manipulation, light massage, hot pack treatment, low-intensity pulsed ultrasound, traction, and exercises. Four studies (Cecchi et al., 2011; Childs et al., 2004; Flynn et al., 2002; Underwood et al., 2007) referred the prescribed SMT technique that clinicians performed in details and others either didn't report or it was up to the clinicians. The number of SMT visits ranged from 2 (Childs et al., 2004; Flynn et al., 2002) to 18 (Vavrek et al., 2015) in 8 studies (Burton et al., 1995; Cecchi et al., 2011; Childs et al., 2004; Flynn et al., 2002; Skargren & Öberg, 1998; Underwood et al., 2007; Vavrek et al., 2015) and was not reported in other studies.

These studies included baseline predictors falling broadly into demographic, history, PROMs, and physical examination domains. The number and variety of predictors examined for intervention outcome in patients with LBP following SMT were substantial as is the lack of consistency across studies. This inconsistency of predictors is likely due to researchers' goals and the availability of data for the secondary analysis studies. There were 2 multivariate linear regression model, 1 modified Poisson regression, and 16 logistic regression models including univariate, bivariate, and multivariate reported. Other relevant factors that have not been yet included in studies so far may

still be associated with the prediction of intervention outcomes. For example, some authors have suggested that an immediate increase in lumbar multifidus recruitment (Fritz et al., 2011; Koppenhaver et al., 2011) or an immediate decrease in spinal stiffness at L3 segment following SMT is associated with clinical success (Fritz et al., 2011).

Overall, the findings suggest that the likelihood of not responding to SMT can be predicted using clinical prediction models, however, there was no trend of improved predictive accuracy in those who would actually respond to SMT. Healthcare clinicians are encouraged to consider integrating the predicted probabilities of recovery into their practice to share decision making with patients and as a method of early treatment and/or referral (Kongsted et al., 2016). For example, a patient with a low probability of recovery with SMT would be a better case to receive additional interventions considering the costs and time involved. Although it does not necessarily result in improved outcomes, more intensive intervention is recommended to patients with a poorer prognosis (da Silva et al., 2017; Hill et al., 2011). Prioritizing consistency in studies addressing clinical prediction models in individuals with LBP may improve model performance in future research.

2.4 References

Abbott, J. H., Flynn, T. W., Fritz, J. M., Hing, W. A., Reid, D. & Whitman, J. M. (2009). Manual physical assessment of spinal segmental motion: Intent and validity. *Manual Therapy*, *14*(1), 36–44. https://doi.org/10.1016/j.math.2007.09.011

Abbott, J. H., McCane, B., Herbison, P., Moginie, G., Chapple, C. & Hogarty, T. (2005). Lumbar segmental instability: A criterion-related validity study of manual therapy assessment. *BMC Musculoskeletal Disorders*, *6*, 1–10. https://doi.org/10.1186/1471-2474-6-56

Angarita-Fonseca, A., Trask, C., Shah, T. & Bath, B. (2019). Stable prevalence of chronic back disorders across gender, age, residence, and physical activity in Canadian adults from 2007 to 2014. *BMC Public Health*, *19*(1), 1–11. https://doi.org/10.1186/s12889-019-7395-8

Axén, I., Jones, J. J., Rosenbaum, A., Lövgren, P. W., Halasz, L., Larsen, K. & Leboeuf-Yde, C. (2005). The nordic back pain subpopulation program: Validation and improvement of a predictive model for treatment outcome in patients with low back pain receiving chiropractic treatment. *Journal of Manipulative and Physiological Therapeutics*, *28*(6), 381–385. https://doi.org/10.1016/j.jmpt.2005.06.008

Axen, I., Rosenbaum, A., Robech, R., Larsen, K. & Leboeuf-Yde, C. (2005). The Nordic back pain subpopulation program: Can patient reactions to the first chiropractic treatment predict early favorable treatment outcome in nonpersistent low back pain? *Journal of Manipulative and Physiological Therapeutics*, *28*(3), 153–158. https://doi.org/10.1016/j.jmpt.2005.02.007

Bath, B., Trask, C., McCrosky, J. & Lawson, J. (2014). A Biopsychosocial Profile of Adult Canadians with and without Chronic Back Disorders: A Population-Based Analysis of the 2009-2010 Canadian Community Health Surveys. *BioMed Research International*, *2014*, 919621. https://doi.org/10.1155/2014/919621

Baumgart, F. & Cordey, J. (2001). Stiffness - An unknown world of mechanical science? *Injury*, *32*, 14–23. https://doi.org/10.1016/S0020-1383(01)00057-2

Bone and Joint Canada. Low Back Pain. (2013). http://www.boneandjointcanada.com/?sec_id=485&msid=3

Brown, B. T., Blacke, A., Carroll, V., Graham, P. L., Kawchuk, G., Downie, A. & Swain, M. (2017). The comfort and safety of a novel rolling mechanical indentation device for the measurement of lumbar trunk stiffness in young adults. *Chiropractic and Manual Therapies*, *25*(21). https://doi.org/10.1186/s12998-017-0153-z

Brunarski, D. (1982). Chiropractic biomechanical evaluations: validity in myofascial low back pain. *J Manipulative Physiol Ther*, *5*(4), 155–161.

Burton, A. K., McClune, T. D., Clarke, R. D. & Main, C. J. (2004). Long-term follow-up of patients with low back pain attending for manipulative care: Outcomes and predictors. *Manual Therapy*, *9*(1), 30–35. https://doi.org/10.1016/S1356-689X(03)00052-3

Burton, A. K., Tillotson, K. M., Main, C. J. & Hollis, S. (1995). Psychosocial predictors of outcome in acute and subchronic low back trouble. *Spine*, *20*(6), 722–728. https://doi.org/10.1097/00007632-199503150-00014

Harvey, D. & Byfieid, D. (1991). Preliminary studies with a mechanical model for the evaluation of spinal motion palpation. *Clin Biomech (Bristol, Avon)*, *6*(2), 79–82. https://doi.org/10.1016/0268-0033(91)90003-9. PMID: 23915479.

Canadian Institute for Health Information (CIHI). *Patient-reported outcome measures (PROMs)*. https://www.cihi.ca/en/patient-reported-outcome-measures-proms

Cecchi, F., Negrini, S., Pasquini, G., Paperini, A., Conti, A., Chiti, M., Zaina, F., Macchi, C. & Molino-Lova, R. (2011). Predictors of functional outcome in patients with chronic low back

pain undergoing back school, individual physiotherapy or spinal manipulation. *European Journal of Physical and Rehabilitation Medicine*, *47*(3), 381–390.

Childs, J. D., Fritz, J. M., Flynn, T. W., Irrgang, J. J., Johnson, K. K., Majkowski, G. R. & Delitto, A. (2004). A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: A validation study. *Annals of Internal Medicine*, *141*(12), 920–928. https://doi.org/10.7326/0003-4819-141-12-200412210-00008

Colloca, C. J. & Keller, T. S. (2001). Stiffness and neuromuscular reflex response of the human spine to posteroanterior manipulative thrusts in patients with low back pain. *J Manipulative Physiol Ther*, *24*(8), 489–500. https://doi.org/10.1067/mmt.2001.118209

Colloca, C. J. & Keller, T. S. (2004). Active trunk extensor contributions to dynamic posteroanterior lumbar spinal stiffness. *Journal of Manipulative and Physiological Therapeutics*, *27*(4), 229–237. https://doi.org/10.1016/j.jmpt.2003.11.001

Colloca, C. J., Keller, T. S., Peterson, T. K. & Seltzer, D. E. (2003). Comparison of dynamic posteroanterior spinal stiffness to plain film radiographic images of lumbar dise height. *Journal of Manipulative and Physiological Therapeutics*, *26*(4), 233–241. https://doi.org/10.1016/S0161-4754(03)00005-8

Cruz, E. B., Canhão, H., Fernandes, R., Caeiro, C., Branco, J. C., Rodrigues, A. M., Pimentel-Santos, F., Gomes, L. A., Paiva, S., Pinto, I., Moniz, R. & Nunes, C. (2020). Prognostic indicators for poor outcomes in low back pain patients consulted in primary care. *PLOS ONE*, *15*(3), 1–15. https://doi.org/10.1371/journal.pone.0229265

da Silva, T., Macaskill, P., Mills, K., Maher, C., Williams, C., Lin, C. & Hancock, M. J. (2017). Predicting recovery in patients with acute low back pain: A Clinical Prediction Model. *European Journal of Pain (United Kingdom)*, *21*(4), 716–726.

https://doi.org/10.1002/ejp.976

Dagenais, S., Caro, J. & Haldeman, S. (2008). A systematic review of low back pain cost of illness studies in the United States and internationally. *Spine Journal*, *8*(1), 8–20. https://doi.org/10.1016/j.spinee.2007.10.005

Eklund, A., Bergström, G., Bodin, L. & Axén, I. (2016). Do psychological and behavioral factors classified by the West Haven-Yale Multidimensional Pain Inventory (Swedish version) predict the early clinical course of low back pain in patients receiving chiropractic care? *BMC Musculoskeletal Disorders*, *17*(75). https://doi.org/10.1186/s12891-016-0933-y

Eklund, A., De Carvalho, D., Pagé, I., Wong, A., Johansson, M. S., Pohlman, K. A., Hartvigsen, J. & Swain, M. (2019). Expectations influence treatment outcomes in patients with low back pain. A secondary analysis of data from a randomized clinical trial. *European Journal of Pain (United Kingdom)*, *23*(7), 1378–1389. https://doi.org/10.1002/ejp.1407

Fadyl, J. & McPherson, K. (2008). Return to work after injury: A review of evidence regarding expectations and injury perceptions, and their influence on outcome. *Journal of Occupational Rehabilitation*, *18*(4), 362–374. https://doi.org/10.1007/s10926-008-9153-0

Ferrari, S., Vanti, C., Pellizzer, M., Dozza, L., Monticone, M. & Pillastrini, P. (2019). Is there a relationship between self-efficacy, disability, pain and sociodemographic characteristics in chronic low back pain? A multicenter retrospective analysis. *Archives of Physiotherapy*, *9*(1), 1–9. https://doi.org/10.1186/s40945-019-0061-8

Ferreira, M. L., Ferreira, P. H., Latimer, J., Herbert, R. D., Maher, C. & Refshauge, K. (2009). Relationship between spinal stiffness and outcome in patients with chronic low back pain. *Clinical Biomechanics*, *14*, 61–67. https://doi.org/10.1016/j.clinbiomech.2016.03.006

Field, J. R. & Newell, D. (2012). Relationship between STarT Back Screening Tool and prognosis

for low back pain patients receiving spinal manipulative therapy. *Chiropractic and Manual Therapies*, *20*(17). https://doi.org/10.1186/2045-709X-20-17

Field, J. R., Newell, D. & McCarthy, P. W. (2010). Preliminary study into the components of the fear-avoidance model of LBP: Change after an initial chiropractic visit and influence on outcome. *Chiropractic and Osteopathy*, *18*, 1–9. https://doi.org/10.1186/1746-1340-18-21

Flynn, T., Fritz, J., Whitman, J., Wainner, R., Magel, J., Rendeiro, D., Butler, B., Garber, M. & Allison, S. (2002). A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine*, *27*(24), 2835–2843. https://doi.org/10.1097/00007632-200212150-00021

Fritz, J. ., Koppenhaver, S. L., Kawchuk, G. N., Teyhen, D. S., Hebert, J. J. & Childs, J. D. (2011). Preliminary investigation of the mechanisms underlying the effects of manipulation: exploration of a multivariate model including spinal stiffness, multifidus recruitment, and clinical findings. *Spine*, *36*(21), 1772–1781. https://doi.org/10.1097/BRS.0b013e318216337d

Fritz, J. M., Piva, S. R. & Childs, J. D. (2005). Accuracy of the clinical examination to predict radiographic instability of the lumbar spine. *European Spine Journal*, *14*(8), 743–750. https://doi.org/10.1007/s00586-004-0803-4

George, S. Z., Lentz, T. A., Beneciuk, J. M., Bhavsar, N. A., Mundt, J. M. & Boissoneault, J. (2020). Framework for improving outcome prediction for acute to chronic low back pain transitions. *PAIN Reports*, *5*(2), e809. https://doi.org/10.1097/pr9.0000000000000809

Geurts, J. W., Willems, P. C., Kallewaard, J. W., Van Kleef, M. & Dirksen, C. (2018). The Impact of Chronic Discogenic Low Back Pain: Costs and Patients' Burden. *Pain Research and Management*, *2018*. https://doi.org/10.1155/2018/4696180

Gross, D. P., Ferrari, R., Russell, A. S., Battié, M. C., Schopflocher, D., Hu, R. W., Waddell, G. & Buchbinder, R. (2006). A population-based survey of back pain beliefs in Canada. *Spine*, *31*(18), 2142–2145. https://doi.org/10.1097/01.brs.0000231771.14965.e4

Harsted, S., Nyirö, L., Downie, A., Kawchuk, G. N., O'Neill, S., Holm, L. & Nim, C. G. (2021). Posterior to anterior spinal stiffness measured in a sample of 127 secondary care low back pain patients. *Clinical Biomechanics*, 105408. https://doi.org/10.1016/j.clinbiomech.2021.105408

Hayden, J. A., Wilson, M. N., Riley, R. D., Iles, R., Pincus, T. & Ogilvie, R. (2019). Individual recovery expectations and prognosis of outcomes in non-specific low back pain: prognostic factor review. *Cochrane Database of Systematic Reviews*, *2019*(11). https://doi.org/10.1002/14651858.CD011284.pub2

Hestbœk, L. & Leboeuf-Yde, C. (2000). Are chiropractic tests for the lumbo-pelvic spine reliable and valid? A systematic critical literature review. *Journal of Manipulative and Physiological Therapeutics*, *23*(4), 258–275. https://doi.org/10.1067/mmt.2000.106097

Hill, J. C., Whitehurst, D. G. T., Lewis, M., Bryan, S., Dunn, K. M., Foster, N. E., Konstantinou, K., Main, C. J., Mason, E., Somerville, S., Sowden, G., Vohora, K. & Hay, E. M. (2011). Comparison of stratified primary care management for low back pain with current best practice (STarT Back): A randomised controlled trial. *The Lancet*, *378*(9802), 1560–1571. https://doi.org/10.1016/S0140-6736(11)60937-9

Hoy, D., Bain, C., Williams, G., March, L., Brooks, P., Blyth, F., Woolf, A., Vos, T. & Buchbinder, R. (2012). A systematic review of the global prevalence of low back pain. *Arthritis and Rheumatism*, *64*(6), 2028–2037. https://doi.org/10.1002/art.34347

Huijbregts, P. A. (2002). Spinal motion palpation: A review of reliability studies. *Journal of*

Manual and Manipulative Therapy, *10*(1), 24–39. https://doi.org/10.1179/106698102792209585

Hurwitz, E. L., Randhawa, K., Yu, H., Côté, P. & Haldeman, S. (2018). The Global Spine Care Initiative: a summary of the global burden of low back and neck pain studies. *European Spine Journal*, *27*(s6), 796–801. https://doi.org/10.1007/s00586-017-5432-9

Iben, A., Rosenbaum, A., Röbech, R., Wren, T. & Leboeuf-Yde, C. (2002). Can patient reactions to the first chiropractic treatment predict early favorable treatment outcome in persistent low back pain? *J Manipulative Physiol Ther 2002, 25*(7), 450–454. https://doi.org/10.1067/mmt.2002.126473.

Iles, R. A., Davidson, M., Taylor, N. F. & O'Halloran, P. (2009). Systematic review of the ability of recovery expectations to predict outcomes in non-chronic non-specific low back pain. *Journal of Occupational Rehabilitation*, *19*(1), 25–40. https://doi.org/10.1007/s10926-008-9161-0

Jenson, K. ., Gemmell, H. & Thiel, H. W. (1993). Motion palpation accuracy using a mechanical spinal model. *European Journal of Chiropractic*, *41*, 67–73.

Kassebaum, N. J., Arora, M., Barber, R. M., Brown, J., Carter, A., Casey, D. C., Charlson, F. J., Coates, M. M., Coggeshall, M., Cornaby, L., Dandona, L., Dicker, D. J., Erskine, H. E., Ferrari, A. J., Fitzmaurice, C., Foreman, K., Forouzanfar, M. H., Fullman, N., Goldberg, E. M., … Zuhlke, L. J. (2016). GBD 2015 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, *388*(10053), 1603–1658. https://doi.org/10.1016/S0140-6736(16)31460-X

Katz, J. N. (2006). Lumbar disc disorders and low-back pain: Socioeconomic factors and consequences. *Journal of Bone and Joint Surgery - Series A*, *88*(SUPPL. 2), 21–24. https://doi.org/10.2106/JBJS.E.01273

Kay, C., Wozniak, E. & Bernstein, J. (2017). Utilization of Health Care Services and Ambulatory Resources Associated with Chronic Noncancer Pain. *Pain Medicine (Malden, Mass.)*, *18*(7), 1236–1246. https://doi.org/10.1093/pm/pnw336

Keller, T. S., Colloca, C. J. & Fuhr, A. W. (1999). Validation of the force and frequency characteristics of the Activator Adjusting Instrument: Effectiveness as a mechanical impedance measurement tool. *Journal of Manipulative and Physiological Therapeutics*, *22*(2), 75–86. https://doi.org/10.1016/S0161-4754(99)70111-9

Kenna, C. & Murtagh, J. (1997). *Back Pain and Spinal Manipulation: A Practical Guid* (2nd ed.). Butterworth-Heinemann.

Kim, L. H., Vail, D., Azad, T. D., Bentley, J. P., Zhang, Y., Ho, A. L., Fatemi, P., Feng, A., Varshneya, K., Desai, M., Veeravagu, A. & Ratliff, J. K. (2019). Expenditures and Health Care Utilization Among Adults With Newly Diagnosed Low Back and Lower Extremity Pain. *JAMA Network Open*, *2*(5), e193676. https://doi.org/10.1001/jamanetworkopen.2019.3676

Kongsted, A., Andersen, C. H., Hansen, M. M. & Hestbaek, L. (2016). Prediction of outcome in patients with low back pain - A prospective cohort study comparing clinicians' predictions with those of the Start Back Tool. *Manual Therapy*, *21*, 120–127. https://doi.org/10.1016/j.math.2015.06.008

Koppenhaver, S. L., Fritz, J. M., Hebert, J. J., Kawchuk, G. N., Childs, J. D., Parent, E. C., Gill, N. W. & Teyhen, D. S. (2011). Association between changes in abdominal and lumbar multifidus muscle thickness and clinical improvement after spinal manipulation. *Journal of*

Orthopaedic and Sports Physical Therapy, *41*(6), 389–399. https://doi.org/10.2519/jospt.2011.3632

Koppenhaver, S. L., Hebert, J. J., Kawchuk, G. N., Childs, J. D., Teyhen, D. S., Croy, T. & Fritz, J. M. (2014). Criterion validity of manual assessment of spinal stiffness. *Manual Therapy*, *19*(6), 589–594. https://doi.org/10.1016/j.math.2014.06.001

Landel, R., Kulig, K., Fredericson, M., Li, B. & Powers, C. M. (2008). Intertester reliability and validity of motion assessments during lumbar spine accessory motion testing. *Phys Ther*, *88*(1), 43–49. https://doi.org/10.2522/ptj.20060179

Langworthy, J. M. & Breen, A. C. (2007). Psychosocial factors and their predictive value in chiropractic patients with low back pain: A prospective inception cohort study. *Chiropractic and Osteopathy*, *15*(5), 1–7. https://doi.org/10.1186/1746-1340-15-5

Latimer, J., Goodsel, M. M., Lee, M., Maher, C. G., Wilkinson, B. N. & Moran, C. C. (1996). Evaluation of a new device for measuring responses to posteroanterior forces in a patient population, part 1: Reliability testing. *Physical Therapy*, *76*(2), 158–165. https://doi.org/10.1093/ptj/76.2.158

Latimer, J., Lee, M., Adams, R. & Moran, C. (1996). An Investigation of the Relationship between Low Back Pain and Lumbar Posteroanterior Stiffness. *Journal of Manipulative & Physiological Therapeutics*, *19*(9), 587–591.

Leboeuf-Yde, C., Axén, I., Jones, J. J., Rosenbaum, A., Løvgren, P. W., Halasz, L. & Larsen, K. (2005). The Nordic Back Pain Subpopulation Program: The long-term outcome pattern in patients with low back pain treated by chiropractors in Sweden. *Journal of Manipulative and Physiological Therapeutics*, *28*(7), 472–478. https://doi.org/10.1016/j.jmpt.2005.07.003

Leboeuf-Yde, C., Grønstvedt, A., Borge, J. A., Lothe, J., Magnesen, E., Nilsson, Ø., Røsok, G.,

Stig, L. C. & Larsen, K. (2004). The nordic back pain subpopulation program: Demographic and clinical predictors for outcome in patients receiving chiropractic treatment for persistent low-back pain. *Journal of Manipulative and Physiological Therapeutics*, *27*(8), 493–502. https://doi.org/10.1016/j.jmpt.2004.08.001

Leboeuf-yde, C. & Larsen, K. (2005). The Nordic back pain subpopulation program: A I-year prospective multicenter study of outcome of presistent low-back pain in chiropractic patients. *Journal of Manipulative & Physiological Therapeutics*, *28*(2), 90–96. https://doi.org/10.1016/j.jmpt.2005.01.0lO

Leboeuf-Yde, C., Rosenbaum, A., Axén, I., Lövgren, P. W., Jørgensen, K., Halasz, L., Eklund, A. & Wedderkopp, N. (2009). The Nordic Subpopulation Research Programme: Prediction of treatment outcome in patients with low back pain treated by chiropractors - does the psychological profile matter? *Chiropractic and Osteopathy*, *17*, 1–13. https://doi.org/10.1186/1746-1340-17-14

Lederer, A.-K., Maly, C., Weinert, T. & Huber, R. (2019). Tissue Stiffness is Not Related to Pain Experience: An Individually Controlled Study in Patients with Chronic Neck and Back Pain. *Evid Based Complement Alternat Med*, *2019*. https://doi.org/10.1155/2019/1907168

Lim, K. L., Jacobs, P. & Klarenbach, S. (2006). A population-based analysis of healthcare utilization of persons with back disorders: Results from the Canadian Community Health Survey 2000-2001. *Spine*, *31*(2), 212–218. https://doi.org/10.1097/01.brs.0000194773.10461.9f

Maher, C., Underwood, M. & Buchbinder, R. (2017). Non-specific low back pain. *The Lancet*, *389*(10070), 736–747. https://doi.org/10.1016/S0140-6736(16)30970-9

Maitland, G., Hengeveld, E., Banks, K. & English, K. (2005). *Maitland's vertebral manipulation*

(7th ed.). Philadelphia: Elsevier.

Malmqvist, S., Leboeuf-Yde, C., Ahola, T., Andersson, O., Ekström, K., Pekkarinen, H., Turpeinen, M. & Wedderkopp, N. (2008). The Nordic back pain subpopulation program: Predicting outcome among chiropractic patients in Finland. *Chiropractic and Osteopathy*, *16*, 1–12. https://doi.org/10.1186/1746-1340-16-13

Moulin, D. E., Clark, A. J., Speechley, M. & Morley-Forster, P. K. (2002). Chronic pain in Canada--prevalence, treatment, impact and the role of opioid analgesia. *Pain Research & Management*, *7*(4), 179–184. https://doi.org/10.1155/2002/323085

Newell, D. & Field, J. (2007). Who will get better? Predicting clinical outcomes in a chiropractic practice. *Clinical Chiropractic*, *10*(4), 179–186. https://doi.org/10.1016/j.clch.2007.10.003

Niemistö, L., Sarna, S., Lahtinen-Suopanki, T., Lindgren, K. A. & Hurri, H. (2004). Predictive factors for 1-year outcome of chronic low back pain following manipulation, stabilizing exercises, and physician consultation or physician consultation alone. *Journal of Rehabilitation Medicine*, *36*(3), 104–109. https://doi.org/10.1080/16501970310019151

Nordin, M., Carragee, E. J., Hogg-Johnson, S., Weiner, S. S., Hurwitz, E. L., Peloso, P. M., Guzman, J., van der Velde, G., Carroll, L. J., Holm, L. W., Côté, P., Cassidy, J. D. & Haldeman, S. (2009). Assessment of Neck Pain and Its Associated Disorders. Results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. *Journal of Manipulative and Physiological Therapeutics*, *32*(2 SUPPL.), S117–S140. https://doi.org/10.1016/j.jmpt.2008.11.016

Owens, E. F., DeVocht, J. W., Gudavalli, M. R., Wilder, D. G. & Meeker, W. C. (2007). Comparison of posteroanterior spinal stiffness measures to clinical and demographic findings at baseline in patients enrolled in a clinical study of spinal manipulation for low back pain.

Journal of Manipulative and Physiological Therapeutics, *30*(7), 493–500. https://doi.org/10.1016/j.jmpt.2007.07.009

Owens Jr, E. F., DeVocht, J. W., Wilder, D. G., Gudavalli, M. R. & Meeker, W. C. (2007). The reliability of a posterior to anterior spinal stiffness measuring system in a population of patients with low back pain. *Journal of Manipulative & Physiological Therapeutics*, *30*(2), 116–123. https://doi.org/10.1016/j.jmpt.2006.12.006

Peterson, C. K., Bolton, J. & Humphreys, B. K. (2012). Predictors of improvement in patients with acute and chronic low back pain undergoing chiropractic treatment. *Journal of Manipulative and Physiological Therapeutics*, *35*(7), 525–533. https://doi.org/10.1016/j.jmpt.2012.06.003

Rapoport, J., Jacobs, P., Bell, N. & Klarenbach, S. (2004). Refining the measurement of the economic burden of chronic diseases in Canada. *Chronic Dis Can*, *25*, 13–21.

Schopflocher, D. (2003). Chronic pain in Alberta: A portrait from the 1996 National Population Health Survey and the 2001 Canadian Community Health Survey. *Edmonton, Alberta: Alberta Health and Wellness*.

Schopflocher, D., Taenzer, P. & Jovey, R. (2011). The prevalence of chronic pain in Canada. *Pain Research & Management : The Journal of the Canadian Pain Society = Journal de La Société Canadienne Pour Le Traitement de La Douleur*, *16*(6), 445–450. https://doi.org/10.1155/2011/876306

Seffinger, M. A., Najm, W. I., Mishra, S. I., Adams, A., Dickerson, V. M., Murphy, L. S. & Reinsch, S. (2004). Reliability of Spinal Palpation for Diagnosis of Back and Neck Pain: a systematic review of the literature. *Spine*, *29*(19), E413–E425. https://doi.org/10.1097/01.brs.0000141178.98157.8e

Skargren, E. I. & Öberg, B. E. (1998). Predictive factors for 1-year outcome of low-back and neck

pain in patients treated in primary care: Comparison between the treatment strategies chiropractic and physiotherapy. *Pain*, *77*(2), 201–207. https://doi.org/10.1016/S0304-3959(98)00101-8

Snodgrass, S. J., Haskins, R. & Rivett, D. A. (2012). A structured review of spinal stiffness as a kinesiological outcome of manipulation: its measurement and utility in diagnosis, prognosis and treatment decision-making. *J Electromyogr Kinesiol*, *22*(5), 708–723. https://doi.org/10.1186/s12891-019-2543-y

Stanton, T. R. & Kawchuk, G. N. (2009). Reliability of assisted indentation in measuring lumbar spinal stiffness. *Manual Therapy*, *14*(2), 197–205. https://doi.org/10.1016/j.math.2008.01.011

Stanton, T. R., Moseley, G. L., Wong, A. Y. L. & Kawchuk, G. N. (2017). Feeling stiffness in the back: A protective perceptual inference in chronic back pain. *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-09429-1

Stochkendahl, M. J., Christensen, H. W., Hartvigsen, J., Vach, W., Haas, M., Hestbaek, L., Adams, A. & Bronfort, G. (2006). Manual examination of the spine: a systematic critical literature review of reproducibility. *Journal of Manipulative and Physiological Therapeutics*, *29*(6), 475-485.e10. https://doi.org/10.1016/j.jmpt.2006.06.011

Thakral, M., Shi, L., Shmerling, R. H., Bean, J. F. & Leveille, S. G. (2014). A stiff price to pay: Does joint stiffness predict disability in an older population? *Journal of the American Geriatrics Society*, *62*(10), 1891–1899. https://doi.org/10.1111/jgs.13070

Tuttle, N., Barrett, R. & Laakso, L. (2008a). Postero-anterior movements of the cervical spine: Repeatability of force displacement curves. *Manual Therapy*, *13*(4), 341–348. https://doi.org/10.1016/j.math.2007.03.005

Tuttle, N., Barrett, R. & Laakso, L. (2008b). Relation between changes in posteroanterior stiffness and active range of movement of the cervical spine following manual therapy treatment. *Spine*, *33*(19), 673–679. https://doi.org/10.1097/BRS.0b013e31817f93f9

Underwood, M. R., Morton, V. & Farrin, A. (2007). Do baseline characteristics predict response to treatment for low back pain? Secondary analysis of the UK BEAM dataset [ISRCTN32683578]. *Rheumatology*, *46*(8), 1297–1302. https://doi.org/10.1093/rheumatology/kem113

Vaillant, M., Pickar, J. G. & Kawchuk, G. N. (2010). Performance and reliability of a variable rate, force/displacement application system. *Journal of Manipulative and Physiological Therapeutics*, *33*(8), 585–593. https://doi.org/10.1016/j.jmpt.2010.08.020

Van Trijffel, E., Anderegg, Q., Bossuyt, P. M. M. & Lucas, C. (2005). Inter-examiner reliability of passive assessment of intervertebral motion in the cervical and lumbar spine: A systematic review. *Manual Therapy*, *10*(4), 256–269. https://doi.org/10.1016/j.math.2005.04.008

Vavrek, D., Haas, M., Neradilek, M. B. & Polissar, N. (2015). Prediction of pain outcomes in a randomized controlled trial of dose-response of spinal manipulation for the care of chronic low back pain. *BMC Musculoskeletal Disorders*, *16*(1), 1–14. https://doi.org/10.1186/s12891-015-0632-0

Vos, T., Allen, C., Arora, M., Barber, R. M., Brown, A., Carter, A., Casey, D. C., Charlson, F. J., Chen, A. Z., Coggeshall, M., Cornaby, L., Dandona, L., Dicker, D. J., Dilegge, T., Erskine, H. E., Ferrari, A. J., Fitzmaurice, C., Fleming, T., Forouzanfar, M. H., … Zuhlke, L. J. (2016). GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Stud. *The Lancet*,

388(10053), 1545–1602. https://doi.org/10.1016/S0140-6736(16)31678-6

Wong, A. Y. L. & Kawchuk, G. N. (2017). The clinical value of assessing lumbar posteroanterior segmental stiffness: a narrative review of manual and instrumented methods. *PM R*, *9*(8), 816–830. https://doi.org/10.1016/j.pmrj.2016.12.001

Wong, A. Y. L., Kawchuk, G., Parent, E. & Prasad, N. (2013). Within- and between-day reliability of spinal stiffness measurements obtained using a computer controlled mechanical indenter in individuals with and without low back pain. *Manual Therapy*, *18*(5), 395–402. https://doi.org/10.1016/j.math.2013.02.003

Wong, A. Y. L., Parent, E. C., Prasad, N., Huang, C., Chan, K. M. & Kawchuk, G. N. (2016). Does experimental low back pain change posteroanterior lumbar spinal stiffness and trunk muscle activity? A randomized crossover study. *Manual Therapy*, *34*, 45–52. https://doi.org/10.1016/j.math.2007.09.013

Xia, T., Long, C. R., Vining, R. D., Gudavalli, M. R., Devocht, J. W., Kawchuk, G. N., Wilder, D. G. & Goertz, C. M. (2017). Association of lumbar spine stiffness and flexion-relaxation phenomenon with patient-reported outcomes in adults with chronic low back pain – a single-arm clinical trial investigating the effects of thrust spinal manipulation. *BMC Complementary and Alternative Medicine*, *17*(1), 303. https://doi.org/10.1186/s12906-017-1821-1

Young, A., Swain, M. S., Kawchuk, G. N., Wong, A. Y. L. & Downie, A. S. (2020). The bench-top accuracy of the VerteTrack spinal stiffness assessment device. *Chiropr Man Ther*, *28*(1), 42. https://doi.org/10.1186/s12998-020-00331-8

Chapter 3. Reliability of a New Loaded Rolling Wheel System for Measuring Spinal Stiffness in Asymptomatic Participants

A version of this chapter has been published. Hadizadeh M., Kawchuk G., Parent E. "Reliability of a new loaded rolling wheel system for measuring spinal stiffness in asymptomatic participants", BMC Musculoskeletal Disorders Journal. 2019; 20(1):176.

Abstract

Background: Few, if any, patient reported symptoms have been shown to be related to objective measures of spine function. Recently, patient-reported measures of disability following spinal manipulative therapy have been associated with an immediate decrease in spinal stiffness obtained by instrumented L3 indentation. Given this novel relation, we anticipate that stiffness measures obtained from locations in addition to L3 may yield valuable information. As such, our research team has developed a new technique to acquire stiffness data continuously over an entire spinal region. The reliability of stiffness measurements obtained by this new technique has yet to be quantified.

Methods: Continuous stiffness testing employs a weighted roller that moves uninterrupted over the spine while measuring the resulting spinal deflection along a subject-specific, laser-defined trajectory. A volunteer sample of asymptomatic participants were assessed in 2 sessions occurring 1 to 4 days apart, with each session scheduled at the same time of day. Each session consisted of 3 trials each beginning at a baseline of ~17N then progressing to a maximally tolerable load as defined from pre-test familiarization trials (~61, 72 or 83 N). Reliability was evaluated with the intraclass correlation coefficient, the standard error of measurement and Bland & Altman analysis.

Results: A total of 17 asymptomatic participants (mean age 29.2 +/- 6 years, 53% female) took part in the study. Overall, the within and between-session reliability of lumbar spine stiffness measures at the maximal tolerable load was excellent ranging from 0.95-1.00 and good to excellent ranging from 0.82-0.93, respectively. Trial averaging was found to reduce standard error of measurement by a mean of 35.2% over all measurement conditions compared to a single trial. Bland and Altman plots for agreement in lumbar spine stiffness measurements varied from -0.3 +/- 1.2 at unloaded condition to -0.4 +/- 2.1 at loaded condition. Data from two participants were removed due to the development of back pain between two sessions.

Conclusion: This study introduced a new technique for measuring spinal stiffness over an entire spinal region in asymptomatic human participants. The new technique produced reliable measurements quantifying the load-displacement values for within-session and between-session assessments.

Keywords: Reliability; Test-Retest; Spine; Stiffness; VerteTrack

3.1 Background

A significant decrease in the mobility of lumbar spine has been reported as a common sign in individuals with low back pain (LBP) (Latimer, Lee, et al., 1996). Previous studies showed that there is a relation between pain and spinal stiffness (Snodgrass et al., 2012). Therefore, spinal stiffness assessment has become a common practice in clinical settings in the management of patients with spine-related pain (Snodgrass et al., 2012; Wong & Kawchuk, 2017). Practitioners routinely evaluate spinal stiffness to provide a basis for diagnosis, prognosis and treatment decision- making (Snodgrass et al., 2012) as well as to monitor the efficacy of treatments such as manipulation (Childs et al., 2004). Typically, the clinical assessment of spinal stiffness involves a manual test where a clinician applies pressure in a posteroanterior (PA) direction to the spinous process of interest (Stanton & Kawchuk, 2009). As stiffness magnitude cannot be quantified precisely with this manual technique, a categorical rating system is often used where the segment of interest is classified as hypomobile, normal, or hypermobile, based on the clinician's perception of stiffness (Stanton & Kawchuk, 2009). Unfortunately, prior studies have shown that clinical judgment of PA testing is highly variable in terms of the magnitude (Latimer et al., 1998), direction (Caling & Lee, 2001) and the speed of applied load (Snodgrass et al., 2012) as well as the discrimination threshold for stiffness perception (Adams, 1995).

Due to low levels of reliability and high variability related to clinical evaluation of spine stiffness, mechanical tools have been developed to quantify the applied loads and tissue displacement that occur during PA testing (Snodgrass et al., 2012; Stanton & Kawchuk, 2009; Wong & Kawchuk, 2017) the majority of which assess force-displacement at a static location. Using this approach, we have shown that patient-reported measures of disability following spinal manipulative therapy

(SMT) are associated with an immediate decrease in spinal stiffness obtained by instrumented L3 indentation (R=0.3) (Fritz et al., 2011; Wong et al., 2015). Given this novel relation, we anticipate that stiffness measures obtained from locations in addition to L3 may yield valuable clinical information. We also hope insights into this area may lead to better management of symptoms of LBP.

As such, our research team has developed a novel device to improve on single-site spinal indentation by employing a loaded rolling wheel system. The reliability of stiffness measurements obtained by this new technique has yet to be quantified. Therefore, the objective of this study was to determine the within- and between-session reliability of lumbar stiffness measurements in asymptomatic participants using this new loaded rolling wheel system (VerteTrack™, VibeDx Corporation, Canada).

3.2 Methods

3.2.1 Participants

A total of 17 consecutive volunteers were recruited using flyers (Appendix A) on campus at University of Alberta. The sample size calculation was based on an estimate used specifically for reliability studies (Walter et al., 1998). Thirteen subjects are needed to detect an ICC of 0.9 with three replications (k=3) against a Null hypothesis of 0.7.

Study participants included asymptomatic males and females between the ages of 18 and 60 with no history of thoracic and lumbar pain within the last 6 months. Participants were excluded from the study if they could not tolerate the stiffness testing procedure, lay prone for 20 min, or had a

history of the following: scoliosis, congenital spinal disorders, prior thoracic or lumbar surgery, spondylolisthesis, cauda equina syndrome, current pregnancy, severe respiratory disease, severe trauma, or a medical 'red flag' such as cancer, spinal infection, fracture, or systemic disease.

3.2.2 Examiner

A research assistant with 6 years of clinical experience in physical therapy and 1 year of experience using the testing device collected all measurements.

3.2.3 Continuous stiffness testing device

The lumbar P-A trunk stiffness was assessed with a mechanical device (Figure 3-1) whose comfort and safety has been studied in a sample of young adults previously (Brown et al., 2017). The device consists of a solid, cube-shaped aluminium frame that provides a rigid support for the roller apparatus. The roller apparatus consists of a vertical rod suspended within a linear bearing to permit near-frictionless vertical translation of two rolling wheels of 70mm diameter with variable inter-wheel spacing (typical 29 mm, ranging from 16 to 54 mm). This inter-wheel spacing adjustments allows the wheels roll over the most prominent part of the paravertebral tissues and not over the spinous processes. This width was obtained for each participant by measuring the distance between the top of the paraspinal tissues using a ruler.

Figure 3-1 Superior view of setting up the device using a laser attached to the wheels.

A stepping motor system (resolution = 0.007 mm) (National Instruments, USA) is used to position the roller along the X (longitudinal, cephald/ caudal), Y (transverse, left-right) axes with built in encoders to confirm motor position. The vertical Z axis employs a stepper motor system (Stepperonline.com, China) that is connected to a cable which raises and lowers the rollers in conjunction with a string potentiometer to quantify vertical position (resolution = 0.020 mm, TE Connectivity, USA). Control of all motors and acquisition of signals is provided by in-house coding using LabVIEW (National Instruments, USA, Figure 3-2). Using this controlling software, it is possible to position the roller in three dimensions. This allows clinicians to manually position the rollers to specific positions along the spine and use a laser pointer mounted on the vertical rod. The laser pointer allows alignment of the rollers to each of the spinous processes of the targeted segments while the device stores the resulting X and Y coordinates. The device then stitches these coordinates together to create a XY trajectory for the wheels to follow. The system then lowers the roller onto the participant and adds additional slack to the Z-axis cable, the roller is free to

move vertically in response to the tissue resistance found along the predefined X-Y trajectory. By repeating this process with additional mass attached to the roller, a continuous measure of the P-A bulk deformation of any spinal region, and hence stiffness, can be quantified.

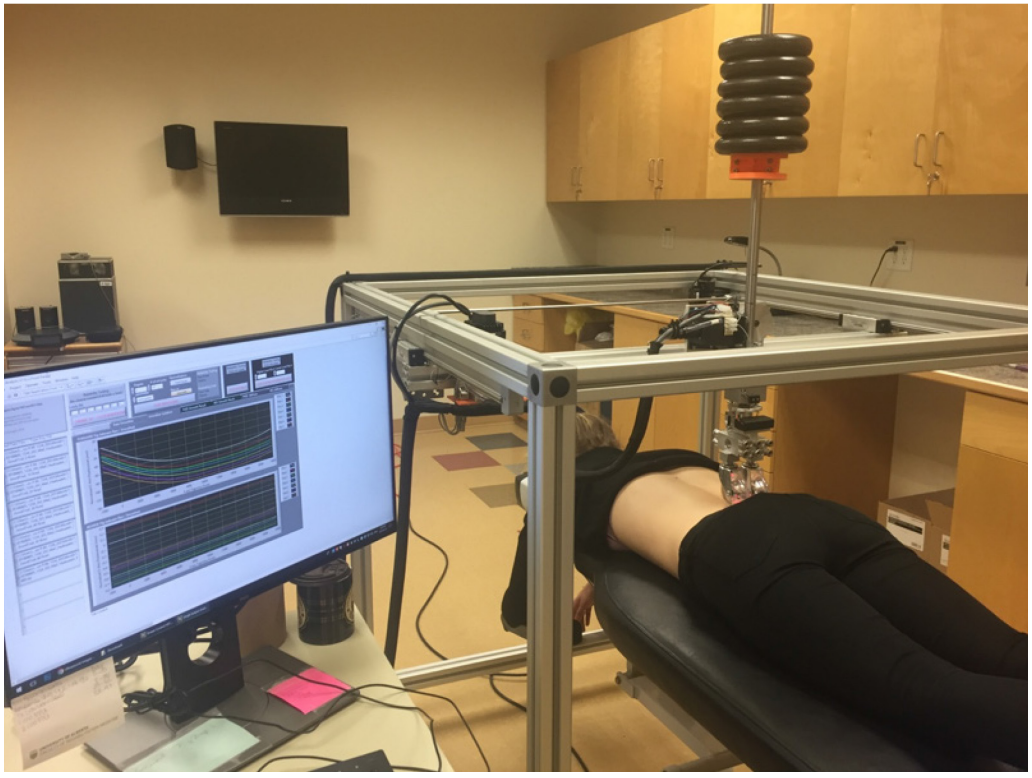


Figure 3-2 Continuous stiffness testing device with participant positioned for the measurement of lumbar spine stiffness and the roller over S1. The device measures displacement which produces by loads applied to the top. The software quantifies stiffness values as a ratio between the applied force and the resultant displacement. The weight of the unloaded roller is 17N. Each additional mass increment is 11N.

3.2.4 Study procedures

Each participant was assessed in 2 separate sessions occurring 1 to 4 days apart. Both sessions were conducted at the same time of day. Prior to testing, consenting participants (Appendix B)

completed self-reported questionnaires on demographics and medical history (Appendix C) as well as completing an 11-point numeric pain rating scales (NPRS-11) before and after each session (Appendix D).

Standardized instructions were given to the participants before testing which included information about how to hold their breath during testing (held expiration), to remain still during testing and to provide feedback if they experienced pain or felt they were resisting the roller wheels. The inter-wheel space of 29mm was used for all participants in both sessions.

To begin using the device, the examiner first manually identified and marked each spinous processes from S1 to T12. The examiner then used the laser system described previously to generate an XY trajectory for the wheels to follow (Figure 3-1). During subsequent stiffness testing, participants were instructed to hold their breath at the end of a normal exhalation for approximately 10s while the device was lowered on to the first trajectory point (S1) and the roller was then automatically moved through the remaining XY trajectory points with the roller free to move vertically in response to spinal topography and tissue resistance. Approximately 10s later, at the last trajectory point (T12), the device was automatically lifted off and returned to the first trajectory point just above S1 while the participant was instructed to continue breathing normally. This process was then repeated with increasing mass attached to the roller with testing ending at either the addition of ~83N in total or when the tolerance of the participant had been reached (pain or muscle contraction) (Figure 3-2). Consistent with previous work (Brown et al., 2017), a rest period of approximately 1 minute was provided between trials.

Prior to data collection, each session began with a familiarization procedure to determine the maximal tolerable load. Participants first experienced the unloaded roller (17N) from S1 to T12. Additional mass was then added in 11N increments until a maximum of ~83N or the maximum tolerable load for each participant was reached.

Following the familiarization procedure, three trials were conducted per session using the unloaded condition and then three additional trials at the maximal tolerable load condition. Data from these trials were used in the reliability analysis. Figure 3-3 shows an example of VerteTrack data output as its rollers move over the back and how the data changes with increased applied loading.

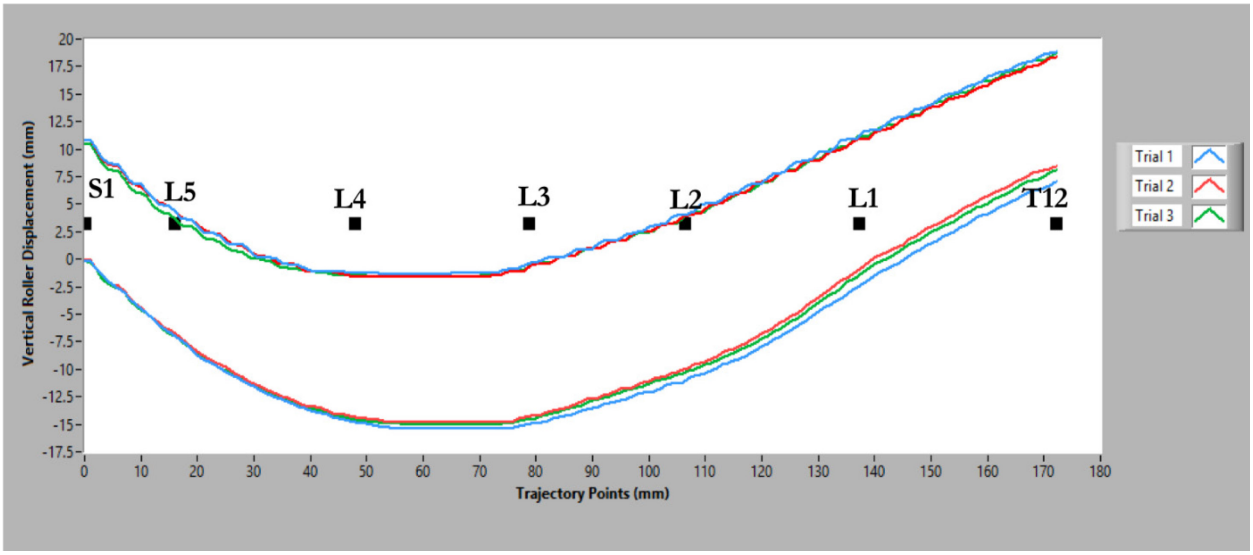


Figure 3-3 An example of VerteTrack data output as its rollers move over the back and how that data changes with increased applied loading. Three trials are shown for the unloaded condition and three for the maximal tolerable load.

In addition, before and during the session, participants were asked to rate any testing-related pain intensity using the NPRS. A reported NPRS of ≥ 2/10 would stop the loading and prior mass

considered as maximum tolerable load (Childs et al., 2005).

These same procedures were repeated in the second session including the familiarization procedure and the reliability tests. All tests were conducted by the same examiner who was blinded to the stiffness assessment results of the first session. Between sessions, participants were asked to 1) maintain their usual physical activities and notice if any new activities had been undertaken between sessions or if new symptoms were present. and 2) to not to wash the spinous process markings on their body so they could be used in the second session.

3.2.5 Data analysis of spinal stiffness

The displacement value for each segment was automatically extracted from a custom program written in LabView and then exported to an Excel file. The roller landing and lifting trajectory points (S1 and T12) of all participants were discarded from the automated extracted data. From the remaining continuous displacement data, stiffness was determined at each of the lumbar spinous process locations with the unloaded roller mass defined the weight of the apparatus (~17 N) and the maximum tolerable load considered as the maximum mass that participants could tolerate with no pain and discomfort (~61, 72 or 83 N) obtained from familiarization process. Stiffness at each spinous process location was then calculated as a ratio between the applied force and the resultant displacement (Fritz et al., 2011).

3.2.6 Statistical analysis

An Intraclass Correlation Coefficient ($ICC_{3, k}$) was calculated to estimate the within session reliability and the between session reliability for stiffness values at each lumbar segment separately. ICC with k indicating 1 provided estimates of the relative reliability for a single trial, and at k= 3 provided estimates of the relative reliability for the average of 3 trials. This model of ICC was chosen because only one examiner was involved in this study, representing a fixed factor for rater (Koo & Li, 2016).

Absolute reliability was obtained by calculating the standard error of measurement (SEM) which is defined as an estimation of the variability expected for observed values when the actual value is held constant (Dudek, 1979). The following formula was used:

SEM= pooled standard deviation $\times \sqrt{}$(1 - ICC)

The Bland and Altman graphs with the difference in spinal stiffness values between session 2 and session 1 (1 minus 2) were plotted against the mean of the 2 test sessions to provide a visual presentation of stiffness values variability (Figure 3-4) (Bland & Altman, 1999). The potential improvement in error when using a single trial, an average of all three trials in determining stiffness was analyzed by comparing the corresponding SEMs.

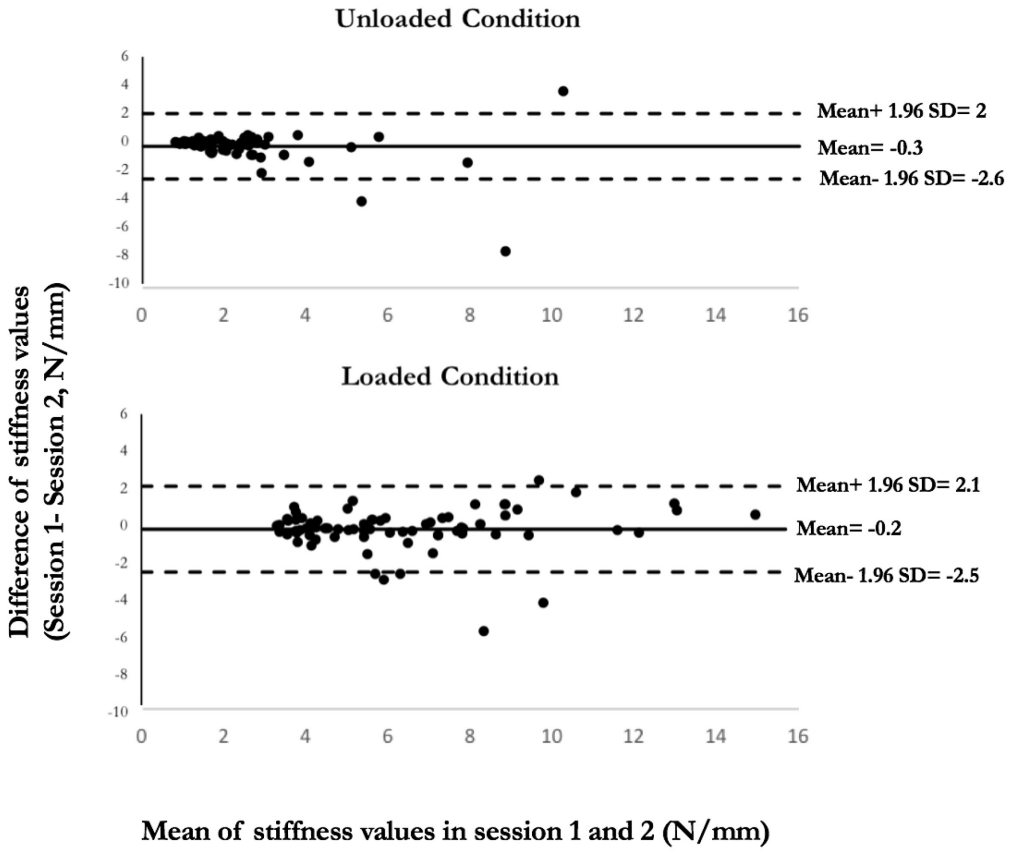


Figure 3-4 Bland-Altman-plot for between- session agreement in spine stiffness measurements. The central horizontal bias reference lines show the average difference between the measurements between the two testing sessions for the unloaded (A) and loaded (B) and the outer lines show the limits of agreement (Bias ± 1.96* standard deviation).

All statistical analyses were performed using IBM SPSS statistics, version 24 (Armonk, New York, USA), (alpha = 0.05). Intraclass Correlation Coefficient values were qualitatively interpreted using the following criteria: 0.00–0.50= poor, 0.50–0.75= moderate, 0.75–0.90= good, and 0.90–1.00= excellent (Koo & Li, 2016).

3.3 Results

Seventeen asymptomatic participants, aged 19-43, fairly homogeneous in terms of age and body mass index were recruited in this study (Table 3-1). No participant was excluded because of not tolerating the testing procedure. As this study was inclusive of asymptomatic participants only, data from two participants were removed from the session 2 due to the development of back pain between the first and second sessions.

Table 3-1 Description of the participants.

Characteristic	**All participants (n=17)**	**Male (n=8)**	**Female (n=9)**
Age (years)	29.2 (6)	29.4 (8.4)	29 (3.1)
Height (cm)	171.3 (14.2)	181.1 (13.4)	162.6 (8)
Weight (kg)	68.5 (15.8)	80.9 (13.3)	57.4 (7.1)
Body mass index (kg/m2)	23.0 (2.4)	24.5 (1.7)	21.7 (2.3)
Numeric pain rating scale (/10)- sessions 1	0.35 (0.7)	0.6 (0.9)	0.1 (0.3)
Numeric pain rating scale (/10)- sessions 2	0.24 (0.4)	0.4 (0.5)	0.1 (0.3)

Values are reported as mean (SD)

The within-session reliability ($ICC_{3,3}$) for the single measures estimated from 0.92 to 1.00 for the unloaded condition and from 0.95 to 1.00 for max tolerable load. In addition, the within-session reliability estimates ($ICC_{3,1}$) for the average of the 3 lumbar spine stiffness measurements ranged from 0.97 to 1.00 for the unloaded condition and from 0.98 to 1.00 for max tolerable load. (Table 3-2). The between-session reliability analysis for the first trial of each session ($ICC_{3,1}$) was ranged from 0.81 to 0.94 for the unloaded condition and from 0.83 to 0.92 for max tolerable load. The between-session reliability estimates of lumbar spine stiffness measurements for the mean of

3 trials ($ICC_{3,1}$) also ranged from 0.75 to 0.96 and 0.82 to 0.93 for unloaded and max tolerable load, respectively (Table 3-2). Overall, the within-session reliability of lumbar spine stiffness measures was excellent and the between-session reliability was good to excellent after removing two participants who reported having back pain.

Table 3-2 Within- session and between- session reliability of stiffness measurements for lumbar tests.

		Minimum Load			Max Tolerable Load		
		Mean (SD) of stiffness values (N/mm)	**$ICC_{3,1}$ (95%CI)/Single measures**	**$ICC_{3,3}$ (95%CI)/ Average measures**	**Mean (SD) of stiffness values (N/mm)**	**$ICC_{3,1}$ (95%CI)/Single measures**	**$ICC_{3,3}$ (95%CI)/Average measures**
Within- session Reliability (session 1/ N=17)	**L5**	2.6 (1.3)	0.99 (0.97_0.99)	1.00 (0.99_1.00)	9.0 (3.4)	0.99 (0.97_0.99)	1.00 (0.99_ 1.00)
	L4	2.0 (1.5)	0.99 (0.97_1.00)	1.00 (0.99_1.00)	6.1 (3.0)	1.00 (0.99_1.00)	1.00 (0.99_ 1.00)
	L3	2.0 (1.9)	0.92 (0.82_0.97)	0.97 (0.93_0.99)	6.2 (5.0)	0.95 (0.89_0.98)	0.98 (0.96_0.99)
	L2	2.2 (1.7)	1.00 (0.99_1.00)	1.00 (0.99_ 1.00)	6.1 (4.4)	1.00 (0.99_1.00)	1.00 (0.99_ 1.00)
	L1	2.7 (2.3)	0.98 (0.95_0.99)	0.99 (0.98_1.00)	7.4 (4.3)	0.99 (0.98_1.00)	1.00 (0.99_1.00)
Within- session Reliability (session 2/ N=15)	**L5**	2.3 (0.8)	0.99 (0.97_0.99)	1.00 (0.99_1.00)	8.1 (2.3)	0.98 (0.96_0.99)	1.00 (0.99_1.00)
	L4	1.6 (0.6)	0.99 (0.98_1.00)	1.00 (0.99_1.00)	5.2 (1.6)	0.99 (0.98_1.00)	1.00 (0.99_1.00)
	L3	1.6 (0.6)	0.99 (0.98_1.00)	1.00 (0.99_ 1.00)	5.0 (1.6)	0.99 (0.98_1.00)	1.00 (0.99_1.00)
	L2	2.4 (2.0)	0.96 (0.90_0.98)	0.99 (0.96_0.99)	5.9 (3.0)	0.99 (0.98_1.00)	1.00 (0.99_1.00)
	L1	3.0 (2.4)	0.96 (0.90_0.99)	0.99 (0.96_1.00)	7.1 (3.4)	0.98 (0.94_0.99)	0.99 (0.98_1.00)
			$ICC_{3,1}$ using the first trial	**$ICC_{3,1}$ using the mean of the 3 trials**		**$ICC_{3,1}$ using the first trial**	**$ICC_{3,1}$ using the mean of the 3 trials**
Between- session Reliability (N=15)	**L5**	2.2 (0.7)	0.81 (0.54_0.93)	0.75 (0.41_0.91)	8.1 (2.3)	0.85 (0.60_0.95)	0.82 (0.55_0.94)
	L4	1.5 (0.5)	0.85 (0.63_0.95)	0.84 (0.60_0.94)	5.2 (1.6)	0.87 (0.66_0.96)	0.93 (0.80_0.98)
	L3	1.5 (0.6)	0.88 (0.67-0.96)	0.86 (0.65_0.95)	5.0 (1.8)	0.86 (0.62_0.95)	0.88 (0.67_0.96)
	L2	2.3 (1.9)	0.94 (0.82_0.98)	0.96 (0.86_0.99)	5.5 (2.6)	0.83 (0.57_0.94)	0.86 (0.63_0.95)
	L1	2.8(2.6)	0.88 (0.68_0.96)	0.75 (0.38_0.92)	6.9 (3.5)	0.92 (0.75_0.97)	0.89 (0.70_0.96)

The effect of averaging a different number of multiple trials on measurement error (standard error of measurements) shows that averaging three repeated measurements reduced the SEM by a mean of 35.2% over all measurement conditions (Table 3-3).

Table 3-3 Changes in standard error of measurement (SEM)

		Single Trial		3 Trials		Mean of 3 Trials (% decrease from 1 measure)	
		Min Load	**Max Load**	**Min Load**	**Max Load**	**Min Load**	**Max Load**
Within-session SEM (N/mm) _ session 1	**L5**	0.2	0.4	0.1	0.2	55.0	40.0
	L4	0.2	0.2	0.1	0.1	50.0	55.0
	L3	1.1	1.6	0.7	0.9	39.1	41.3
	L2	0.3	0.4	0.2	0.2	36.7	52.5
	L1	0.9	0.7	0.5	0.4	42.2	45.7
Within-session SEM (N/mm) _ session 2	**L5**	0.1	0.3	0.1	0.2	30.0	33.3
	L4	0.1	0.2	0.04	0.1	60.0	50.0
	L3	0.1	0.3	0.03	0.2	70.0	50.0
	L2	0.4	1.0	0.2	0.6	40.0	43.0
	L1	0.8	0.4	0.6	0.3	20.0	30.0
Between-session SEM (N/mm)	**L5**	0.3	1.0	0.3	0.7	0.0	30.0
	L4	0.2	0.6	0.2	0.4	0.0	33.3
	L3	0.2	0.7	0.2	0.5	0.0	28.6
	L2	0.4	0.6	0.3	0.6	25.0	0.0
	L1	1.3	1.2	0.9	0.9	30.8	25.0

Abbreviations: SEM, standard error of measurements

3.4 Discussion

In this study, we evaluated the test- retest reliability of spinal stiffness measurements in asymptomatic individuals using a new device that collects continuous measures from all lumbar levels and found excellent within and between-session reliability at the maximal tolerable load. No control group was required for the design of this study.

3.4.1 Within- and between- session reliability

Our within- session reliability values for stiffness measurement are similar to prior data reported by Wong et al. (ICC, 0,99) (Wong et al., 2013b), and comparable to other studies using single point indentation devices (ICC, 0.96 to 0.98) (Edmondston et al., 1998; Latimer, Goodsel, et al., 1996; Shirley et al., 2002). However, the between- session reliability values at the maximal tolerable load for the averaged measurements (0.90 to 0.94) are lower than Wong et al's prior study (Wong et al., 2013b) (0.98) but better than those reported from the previous automated techniques (0.85 and 0.88) (Lee & Svensson, 1990; Shirley et al., 2002). The improved between- session reliability of mechanical indenter in Wong et al.'s study might be attributed to the larger sample size. In addition, while Wong et al. used ultrasound to identify the spinous process location, we used an alternative technique by asking each participant to not wash our spinous process markings on their body for use in the next session. We selected this technique as it is not susceptible to ultrasound operator error between sessions – the same markings are used in each participant for each session. Importantly, even if these marking are incorrect in terms of the spinous processes identified, using the same markings are better suited to this reliability study. Therefore, the between- session reliability will not have been affected by the verification of the spinous process location using traditional manual technique.

The Bland and Altman plots show the majority of observations fall on or very near the mean resulting in a high level of agreement between the two measurement sessions. Any difference in stiffness between sessions may be attributed to the individual differences in the time between sessions or the individual activities of the participants between sessions. The Bland and Altman

plots show less reliability at higher stiffness measurements in both unloaded and loaded conditions. Possible explanations for this observation between sessions may include a variety of patient-based factors such as activity level and apprehension level.

3.4.2 Loaded versus unloaded conditions

The unloaded conditions and the loaded (max tolerable load) conditions did not differ significantly in terms of within-session and between session reliability. This is shown by the ICC confidence intervals presented in Table 3-2 which overlap for most corresponding estimates for the unloaded and loaded conditions. This suggests that the device provided reliable values regardless of the applied load. However, for the majority of the comparisons between the corresponding unloaded and loaded ICC point estimates, when there is a difference, the point estimate of the loaded condition is better. Clinically, the unloaded condition will likely be more tolerable in patients with LBP and our results confirm that the unloaded condition can provide reliable data.

3.4.3 Changes in measurement error by multiple trial

Our study found that using an average of the three trials to create within- session stiffness values showed a reduction in SEMs as compared with a single trial. This is consistent with previous studies (Wong et al., 2013b) that showed using an average of three measurements improved the measurement error. Therefore, we suggest taking the results from an average of 3 trials if possible to calculate the stiffness of a spinal region using VerteTrack.

3.4.4 Limitations and future research

The study protocol, which was designed for a research study on reliability, took 30-45 minutes including the familiarization procedure. Using single trials only, the total time to complete testing is ~ 12 minutes (Young et al., 2020).

While participants returned at similar times on separate sessions, it is currently unclear whether better control of inter-session intervals would improve between- session reliability results; it is impossible to know if a change in reliability in the second session is the result of differences in the participant over time, variability in the measurement process, or both. This is a drawback of reliability testing over multiple days. Furthermore, the measures obtained by a loading device such as this will always be influenced by the viscoelastic properties of the target tissues in their current state. As such, the reliability of this device is dependent on providing adequate recovery time between trials.

While we expect that the reliability of the device may change when used to evaluate spinal pathology, this device may be contraindicated in specific pathologies as well (e.g., fracture, metastatic disease). Further studies are needed to define relative and absolute contraindications for VerteTrack use. It is important to note that the reliability of the VerteTrack is likely decreased by patient-based factors such as voluntary/involuntary muscle contraction, changes in patient position during testing and inconsistent patient breathing procedures. Future identification of these factors and the magnitude of their impact is warranted.

3.5 Conclusions

This study evaluated the reliability of a device capable of measuring spinal stiffness continuously over an entire spinal region in asymptomatic human participants. The new technique was shown to produce reliable measurements in quantifying the load-displacement values for within- session and between- session assessments. The resulting data may have greater clinical utility in that spinal stiffness can be obtained not only at one level, but over the entire spinal region of interest.

List of abbreviations

LBP: Low Back Pain; SEM: Standard Error of Measurement; ICC: Intraclass Correlation Coefficients; NPRS: Numerical Pain Rating Scale.

Declarations

Ethics approval and consent to participate

Written informed consent was obtained from all participants in accordance with protocol number Pro00061205 approved by the University of Alberta Health Research Ethics Board (Appendix E).

Consent for publication

Not applicable.

Availability of data and material

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

No competing interests

Funding

Not funded

Authors' contributions

MH, GK and EP were responsible for the design, conducting the study, data analysis and writing process. All authors read and approved the final manuscript.

Acknowledgments

We would like to thank all the participants who took part in this research for their time and help.

3.6 References

Adams, R. (1995). A psychophysical evaluation of manual stiffness discrimination. *Australian Journal of Physiotherapy*, *41*(3), 161–167. https://doi.org/10.1016/S0004-9514(14)60426-8

Bland, J. M. & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, *8*(2), 135–160. https://doi.org/10.1177/096228029900800204

Brown, B. T., Blacke, A., Carroll, V., Graham, P. L., Kawchuk, G., Downie, A. & Swain, M. (2017). The comfort and safety of a novel rolling mechanical indentation device for the measurement of lumbar trunk stiffness in young adults. *Chiropractic and Manual Therapies*, *25*(21). https://doi.org/10.1186/s12998-017-0153-z

Caling, B. & Lee, M. (2001). Effect of direction of applied mobilization force on the posteroanterior response in the lumbar spine. *Journal of Manipulative & Physiological Therapeutics*, *24*(2), 71–78. https://doi.org/10.1067/mmt.2001.112568

Childs, J. D., Fritz, J. M., Flynn, T. W., Irrgang, J. J., Johnson, K. K., Majkowski, G. R. & Delitto, A. (2004). A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: A validation study. *Annals of Internal Medicine*, *141*(12), 920–928. https://doi.org/10.7326/0003-4819-141-12-200412210-00008

Childs, J. D., Piva, S. R. & Fritz, J. M. (2005). Responsiveness of the Numeric Pain Rating Scale in Patients with Low Back Pain. *Spine*, *30*(1), 1331–1334. https://doi.org/10.1097/01.brs.0000164099.92112.29

Dudek, F. J. (1979). The continuing misinterpretation of the standard error of measurement. *Psychological Bulletin*, *86*(2), 335–337. https://doi.org/10.1037/0033-2909.86.2.335

Edmondston, S. J., Allison, G. T., Gregg, C. D., Purden, S. M., Svansson, G. R. & Watson, a. E.

(1998). Effect of position on the posteroanterior stiffness of the lumbar spine. In *Manual therapy* (Vol. 3, Issue 1, pp. 21–26). https://doi.org/10.1054/math.1998.0312

Fritz, J. M., Koppenhaver, S. L., Kawchuk, G. N., Teyhen, D. S., Hebert, J. J. & Childs, J. D. (2011). Preliminary investigation of the mechanisms underlying the effects of manipulation: exploration of a multivariate model including spinal stiffness, multifidus recruitment, and clinical findings. *Spine*, *36*(21), 1772–1781. https://doi.org/10.1097/BRS.0b013e318216337d

Koo, T. K. & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Latimer, J., Goodsel, M. M., Lee, M., Maher, C. G., Wilkinson, B. N. & Moran, C. C. (1996). Evaluation of a new device for measuring responses to posteroanterior forces in a patient population, Part 1: Reliability testing. *Physical Therapy*, *76*(2), 158–165.

Latimer, J., Lee, M. & Adams, R. D. (1998). The effects of high and low loading forces on measured values of lumbar stiffness. *Journal of Manipulative and Physiological Therapeutics*, *21*(3), 157–163.

Latimer, J., Lee, M., Adams, R. & Moran, C. (1996). An Investigation of the Relationship between Low Back Pain and Lumbar Posteroanterior Stiffness. *Journal of Manipulative & Physiological Therapeutics*, *19*(9), 587–591.

Lee, M. & Svensson, N. L. (1990). Measurement of stiffness during simulated spinal physiotherapy. *Clin Phys Physiol Meas*, *11*(3), 201–207. https://doi.org/10.1088/0143-0815/11/3/002

Shirley, D., Ellis, E. & Lee, M. (2002). The response of posteroanterior lumbar stiffness to repeated

loading. *Manual Therapy*, *7*(1), 19–25. https://doi.org/10.1054/math.2001.0432

Snodgrass, S. J., Haskins, R. & Rivett, D. A. (2012). A structured review of spinal stiffness as a kinesiological outcome of manipulation: its measurement and utility in diagnosis, prognosis and treatment decision-making. *J Electromyogr Kinesiol*, *22*(5), 708–723. https://doi.org/10.1186/s12891-019-2543-y

Stanton, T. R. & Kawchuk, G. N. (2009). Reliability of assisted indentation in measuring lumbar spinal stiffness. *Manual Therapy*, *14*(2), 197–205. https://doi.org/10.1016/j.math.2008.01.011

Walter, S. D., Eliasziw, M. & Donner, A. (1998). *Sample size and optimal designs for reliability studies*. *17*(1), 101–110. https://doi.org/10.1002/(sici)1097-0258(19980115)17:1<101::aid-sim727>3.0.co;2-e

Wong, A. Y. L. & Kawchuk, G. N. (2017). The clinical value of assessing lumbar posteroanterior segmental stiffness: a narrative review of manual and instrumented methods. *PM R*, *9*(8), 816–830. https://doi.org/10.1016/j.pmrj.2016.12.001

Wong, A. Y. L., Kawchuk, G., Parent, E. & Prasad, N. (2013). Within- and between-day reliability of spinal stiffness measurements obtained using a computer controlled mechanical indenter in individuals with and without low back pain. *Manual Therapy*, *18*(5), 395–402. https://doi.org/10.1016/j.math.2013.02.003

Wong, A. Y. L., Parent, E. C., Dhillon, S. S., Prasad, N. & Kawchuk, G. N. (2015). Do participants with low back pain who respond to spinal manipulative therapy differ biomechanically from nonresponders, untreated controls or asymptomatic controls? *Spine*, *40*(17), 1329–1337. https://doi.org/10.1097/BRS.0000000000000981

Young, A., Swain, M. S., Kawchuk, G. N., Wong, A. Y. L. & Downie, A. S. (2020). The bench-

top accuracy of the VerteTrack spinal stiffness assessment device. *Chiropr Man Ther*, *28*(1), 42. https://doi.org/10.1186/s12998-020-00331-8

Chapter 4. A Consensus Approach Toward the Standardization of Spinal Stiffness Measurement Using a Loaded Rolling Wheel Device: Results of a Delphi Study

A version of this chapter has been published. Hadizadeh M., Kawchuk G., French S. "A consensus approach toward the standardization of spinal stiffness measurement using a loaded rolling wheel device: results of a Delphi study", BMC Musculoskeletal Disorders journal, 2021; 22(1):436.

Abstract

Background: Spinal stiffness assessment has the potential to become an important clinical measure. Various spinal stiffness-testing devices are available to help researchers objectively evaluate the spine and patient complaints. One of these is VerteTrack, a device capable of measuring posteroanterior displacement values over an entire spinal region. This study aimed to develop a best-practice protocol for evaluating spinal stiffness in human participants using VerteTrack.

Methods: Twenty-five individuals with research experience in measuring spinal stiffness, or who were trained in spinal stiffness measurement using the VerteTrack device, were invited to participate in this 3-Round Delphi study. Answers to open-ended questions in Round 1 were thematically analyzed and translated into statements about VerteTrack operation for spinal stiffness measurements. Participants then rated their level of agreement with these statements using

a 5-point Likert scale in Rounds 2 and 3. A descriptive statistical analysis was performed. Consensus was achieved when at least 70% of the participants either strongly agreed, agreed, (or strongly disagreed, disagreed) to include a statement in the final protocol.

Results: Twenty participants completed Round 1 (80%). All these participants completed Rounds 2 and 3. In total, the pre-defined consensus threshold was reached for 67.2% (123/183) of statements after three rounds of surveys. From this, a best-practice protocol was created.

Conclusions: Using a Delphi approach, a consensus-based protocol for measuring spinal stiffness using the VerteTrack was developed. This standard protocol will help to improve the accuracy, efficiency, and safety of spinal stiffness measurements, facilitate the training of new operators, increase consistency of these measurements in multicenter studies, and provide the synergy and potential for data comparison between spine studies internationally. Although specific to VerteTrack, the resulting standard protocol could be modified for use with other devices designed to collect spinal stiffness measures.

Keywords: Spinal stiffness, Mechanical instruments, VerteTrack, Spinal pain, Delphi

4.1 Background

Low back pain (LBP) is the most burdensome of musculoskeletal conditions globally affecting ~ 7.5% of the world's population (~ 577 million people) (Wu et al., 2020). For up to 90% of people presenting with LBP, the specific cause of their pain cannot be clearly identified resulting in a label of non-specific LBP (Maher et al., 2017). The current treatment of LBP mainly focuses on pain management while the causes of pain are rarely addressed. Quantitative assessments of the spine and patient complaints related to LBP may help with the identification of causes, improve the management of this condition, and reduce health care system costs.

Advances in science and technology over the past few decades have made several devices available to objectively assess clinical characteristics of patients including spinal stiffness. Stiffness is considered an important spinal biomechanical measure and has long been recognized by both patients and clinicians as one of the characteristic features of the back (Colloca, 2010). Therefore, stiffness has been widely used in the management of patients with back pain for diagnosis, prognosis, clinical decision-making, and the evaluation of manipulative techniques (Snodgrass et al., 2012).

An increase or decrease in spinal stiffness has been found to be related to LBP. Specifically, previous studies have demonstrated that some patients with LBP have abnormal levels of spinal stiffness (Wong & Kawchuk, 2017) and that these patients experience an immediate and sustained decrease in spinal stiffness for 1 week following spinal manipulative therapy (Jun et al., 2020; Wong et al., 2015). Moreover, researchers reported an increase in posteroanterior (PA) stiffness in participants with LBP compared to when participants had little or no pain, while asymptomatic

controls showed insignificant changes in PA stiffness over time (Latimer et al., 1996). A reduction in stiffness has also been shown to be associated with self-reported measures of disability (Fritz et al., 2011; Wong et al., 2015). These findings suggest that restoration of normal spinal stiffness and mobility plays an important role in some patients with LBP by improving spinal function and reducing pain although a casual relation between stiffness and these outcomes has not been confirmed. Therefore, further exploration of spinal stiffness assessment is warranted. While there are various spinal stiffness-testing devices available to objectively evaluate the spinal complaints (Snodgrass et al., 2012; Wong & Kawchuk, 2017), there is no standard operating protocol for spinal stiffness measurement.

Having a standard data collection protocol for spinal stiffness assessment would facilitate comparison of devices and data between studies. Our research team developed a novel device, the VerteTrack, to improve on single-site spinal indentation by employing a loaded rolling wheel system. Several identical devices have been manufactured and are in use in multiple research centers over the past 6 years. In this Delphi study, our goal was to develop a best-practice protocol for evaluating spinal stiffness in human participants using VerteTrack, a spinal stiffness measurement device shown to be safe (Brown et al., 2017), reliable (Hadizadeh et al., 2019), and accurate (Young et al., 2020).

4.2 Methods

This study used a standard Delphi methodology to achieve consensus. The Delphi method is a reliable and structured method of obtaining a consensus of opinion from a group of experts or knowledgeable participants (Hasson et al., 2000) in areas where existing research is limited. The

Delphi method is particularly recommended for areas where controversy, debate, or a lack of clarity exist (Iqbal & Pipon-Young, 2009).

4.2.1 Selection of participants

As our lab manufactured the device in question, we know of all the research centers that possess the device and all the staff who were trained on the device. We contacted these centers and asked them to provide us with an updated contact list of those who were trained/ used the device since their initial training session. Thus, all individuals trained in VerteTrack methods and/or having previous experience using the VerteTrack device were invited to participate in the Delphi process (n = 25 individuals from 9 different institutions in 7 different countries). Potential participants were asked to participate in the study if they were willing to participate, have access to the internet over the course of the study, and were able to commit time to complete the surveys. Written consent was obtained from all participants after being informed about the project by adding a consent question to the start of Round 1.

4.2.2 Delphi-survey procedure

The Delphi survey involved three sequential rounds of deidentified online questionnaires provided over 4 months (Sep-Dec 2020). Study data were collected and managed using REDCap (Harris et al., 2009) electronic data capture tools provided by the Women & Children's Health Research Institute at the University of Alberta. We contacted the research centers that are equipped with the device and asked them to send us the email addresses of those who were trained or collected data using the device. E-mail addresses were then entered into the REDcap website. All potential participants were sent an invitation email to participate in the Delphi process containing a link to

the online survey. Participants were requested to complete each questionnaire within 2 weeks. Two automated e-mail reminders per round were sent out to non-responders at 1 week and the day before the due date. If participants were not able to complete the questionnaires within the 2 weeks, they were provided with additional reminders and extra time to respond. Each survey took 20–30min to complete.

Participants were allowed to save their answers and return to complete the questionnaire over several sessions. Prior to the commencement of this study, consensus was defined when at least 70% of the participants in Rounds 2 and 3 either strongly agreed, agreed, (or strongly disagreed, disagreed) to include a statement in the final protocol. These levels of agreement have been considered appropriate in previous Delphi studies (Akins et al., 2005; Hasson et al., 2000; Keeney et al., 2006, 2011; Wells et al., 2014). Figure 4-1 summarizes the stages of the Delphi method in this study.

In order to improve the structure and readability of questions, the Round 1 questionnaire was first piloted with three colleagues. Based on their feedback, Round 1 questions were revised and finalized (Appendix F). MH and GNK designed the Round 1 of the survey. This round included questions regarding basic demographic information and 21 open-ended questions inquiring about participant recruitment for VerteTrack testing, device safety, instructions given to research participants, and technical issues. This round aimed to review the comprehensiveness and relevance of the items and provide suggestions for the eventual protocol. Items for Round 2 of the survey were generated by comments from the first round that suggested removing, aggregating, or retaining items from the first round (Appendix G).

Only those who completed round 1 were invited to participate in Round 2. In this round, each participant received a survey comprising 171 statements. The goal of this round was to reach consensus on a standard protocol. In Round 2, participants were asked to indicate their anonymous opinion by ranking statements along a five-point Likert scale for agreement ("strongly agree", "agree", "neither agree nor disagree", "disagree", "strongly disagree"). Additionally, a free-text comment section for each question was available for participants to express any further thoughts or opinions. Round 2 also included four new open-ended questions derived from Round 1. Participants were required to rate every single item to be able to move on with the questionnaire.

Round 3 of the study comprised the same list and grading scale as Round 2 with an additional graphical description of findings from the previous round (Appendix H). The graphic information identified the percentage of total respondents that selected each possible score for the given item in Round 2. The respondents, therefore, were given an opportunity to modify or confirm their answers after viewing the scoring results using the same Likert scale from the previous round. The revised and new statements proposed by participants were added in Round 3 yielding a total of 183 statements. Using the consensus results obtained from Round 3, the authors created a written protocol for use of the VerteTrack device in collecting spinal stiffness measures.

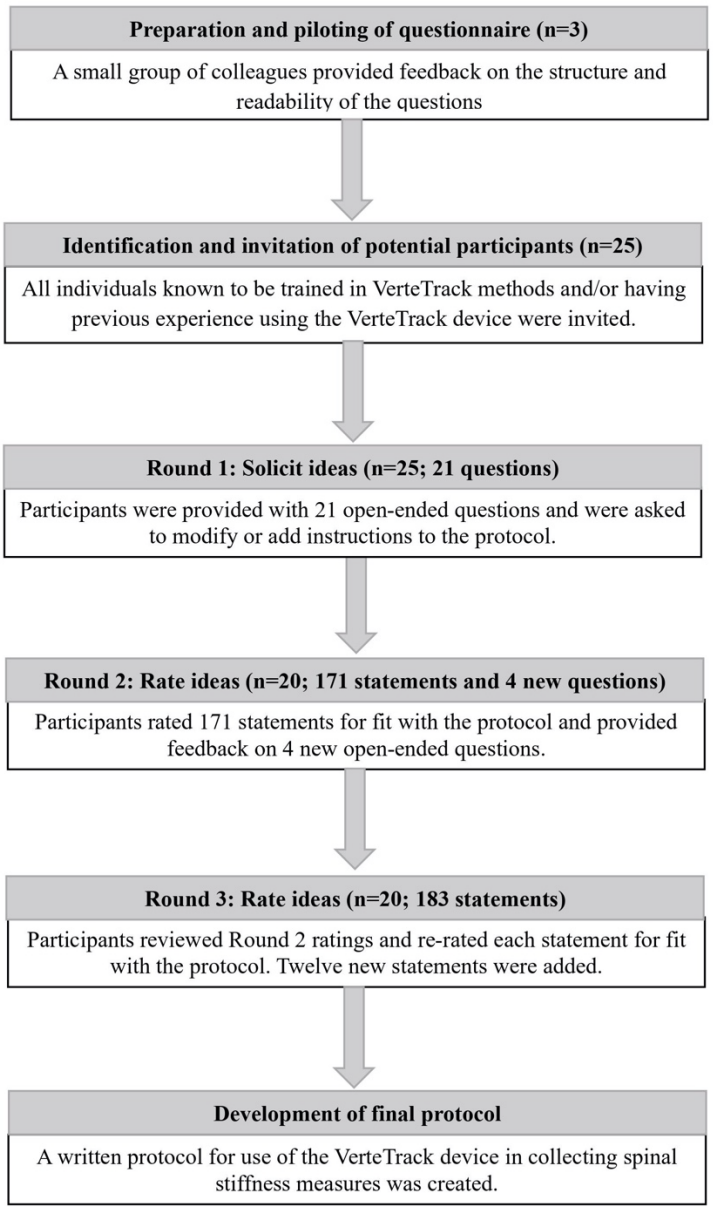


Figure 4-1 Stages of the Delphi technique to standardize spinal stiffness

4.2.3 Analysis

Deidentified data were analyzed by encoding participants with their survey ID numbers. Data from the REDCap tool was downloaded into a Microsoft Excel version 16.45 after each round. Descriptive statistics were used to describe the participants' demographic characteristics.

Responses to open-ended questions in the Round 1 and participants' comments in Round 2 were thematically analyzed with MH and GNK discussing the qualitative responses. MH, GNK and SF met to discuss the items for the consensus statements in Rounds 2 and 3. The quantitative responses from the participants' ratings in Rounds 2 and 3 were analyzed descriptively using medians, ranges, and percentages.

4.3 Results

Of the 25 individuals invited to participate in this Delphi study, 20 participants completed Round 1 (80% response rate), 20/20 completed Round 2 (100.0% response rate), and 20/20 completed Round 3 (100.0% response rate). The reasons for 5/25 participants not responding to the initial invitation email were not identified. Table 4-1 presents the demographic characteristics of participants at baseline.

Participants had different experiences working with the device that ranged from receiving training to perform measurements of spinal stiffness in a population of 180 patients with back pain. In total, the pre-defined consensus threshold was reached for 67.2% (123/183) of statements after three rounds of surveys. Results from Round 3 were presented in Table 4-2. The number of consensus statements under each category was listed in Table 4-3. Items with 70% or more consensus from Round 3 were used to create the best practice protocol for the VerteTrack device (Appendix I).

Table 4-1 Baseline characteristics of Delphi participants (n=20)

Baseline characteristics	Value [a]
Gender (% female)	35
Age (years)	32.5 ± 8.3
Years of clinical experience	5.6 ± 6.6
The country in which the measurement was performed (%).	
Australia	20
Canada	30
Denmark	15
France	5
Honk Kong	10
USA	20
Highest educational qualification (%)	
BSc	15
MSc	45
Ph.D.	35
D.C.	5
Occupation at the time of the study (%)	
Assistant professor	15
Senior lecturer	15
Post-doc fellow	5
Research coordinator	5
Research assistant	10
Student	20
Chiropractor	15
Physiotherapist	15
Primary discipline (%)	
Chiropractic	55
Physiotherapy	25
Other	20
The number of participants assessed using the VerteTrack device (min-max)	0-180

[a] Values are mean ± SD unless otherwise indicated.

Table 4-2 Median value of Likert scale data and agreement level for all statements from Round 3

Domain	Consensus statement	Median (Range)	Percentage of respondents rating each statement		
			Agree (%)	Neutral (%)	Disagree (%)
Inclusion criteria	The ability to tolerate a load of at least 40 N.	1 (1-3)	95.0*	5.0	0.0
	BMI under 40 for ease of palpation.	1 (1-3)	85.0*	15.0	0.0
	18 years or older.	3 (1-4)	30.0	55.0	15.0
	Chronic back pain.	3 (1-5)	25.0	45.0	30.0
Exclusion criteria	Pregnancy.	1 (1-2)	100.0*	0.0	0.0
	Skin lesion, infection, or open wounds over the back region.	1 (1-1)	100.0*	0.0	0.0
	Unable to lie in the prone position (e.g., severe deformities to spine or limbs, static tremor, uncontrolled epilepsy).	1 (1-1)	100.0*	0.0	0.0
	Serious spinal pathology (e.g., spinal tumor, fracture, infectious disorder, osteoporosis, or other bone demineralizing condition).	1 (1-2)	100.0*	0.0	0.0
	Unable to maintain their breathing cycle in passive expiration (functional residual capacity) for at least 10 seconds.	1 (1-2)	100.0*	0.0	0.0
	Unable to follow instructions (e.g., those with dementia or children (age under 18) who may move during the test.	1 (1-3)	95.0*	0.5	0.0
	A head, neck, or thoracoabdominal surgery within the last 6 months.	1 (1-3)	90.0*	10.0	0.0
	Unstable spondylolisthesis.	1 (1-5)	85.0*	10.0	5.0
	Unstable and/or acute disc herniation or injury.	1.5 (1-4)	75.0*	20.0	5.0
	People who do not feel comfortable with the VerteTrack procedure.	1 (1-4)	75.0*	20.0	5.0
	Unstable heart condition.	2 (1-4)	70.0*	15.0	15.0
	Claustrophobia (a fear of being in closed or small spaces).	2 (1-5)	65.0	15.0	20.0
	Acute pain in the test area (depends on whether a participant can tolerate the loading and how long the aggravated pain will subside).	2 (1-5)	60.0	20.0	20.0
	Obesity using BMI (e.g., BMI>30).	2 (1-4)	55.0	15.0	30.0
	Hyperalgesia (an abnormally increased sensitivity to pain).		55.0	25.0	20.0
	Obesity using waist circumference (e.g., waist circumference more than 35 inches in women).	2.5 (1-5)	50.0	15.0	35.0
	Previous sacrum trauma/sensitive sacrum.	3 (1-5)	45.0	25.0	30.0

	Spinal canal stenosis.	3 (1-5)	35.0	25.0	40.0
	Participants with exaggerated spinal curves e.g., thoracic hyper-kyphosis.	3 (1-5)	30.0	25.0	45.0
	People with asthma, colds, or breathing disorders.	4 (1-5)	25.0	20.0	55.0
	History of spine surgery (depends on whether a participant can tolerate the loading and how long the aggravated pain will subside).	3.5 (1-5)	25.0	25.0	50.0
	Scoliosis.	3.5 (1-5)	20.0	30.0	50.0
	Tenderness in the test area (depends on whether a participant can tolerate the loading and how long the aggravated pain will subside).	4 (1-5)	15.0	20.0	65.0
Pregnancy	A pregnant woman should not participate at any stage of pregnancy.	3 (1-5)	45.0	30.0	25.0
	From the first day of pregnancy to 3 months postpartum.	3 (1-5)	45.0	35.0	20.0
	Excluded from the second trimester.	3 (1-5)	40.0	35.0	25.0
	From confirmation of pregnancy till 6 weeks postpartum.	3 (1-5)	35.0	30.0	35.0
	From the first day of pregnancy till 1 month postpartum.	3 (1-5)	30.0	40.0	30.0
	From the first day of pregnancy to the day following the delivery.	3 (1-5)	25.0	35.0	40.0
	From confirmation of pregnancy to 12 months postpartum.	4 (1-5)	15.0	30.0	55.0
Participants' familiarization procedures	Remind the participants once again some points to note e.g., hold breath during the measurement.	1 (1-2)	100.0*	0.0	0.0
	Make sure participants have understood the procedure and don't have any questions.	1 (1-2)	100.0*	0.0	0.0
	Practice breathing protocol with the participant before beginning the measurements.	1 (1-2)	100.0*	0.0	0.0
	Some reassurance that while they may feel pressure on the spine, the device will not cause any harm.	1 (1-2)	100.0*	0.0	0.0
	Explain that there is an emergency stop.	1 (1-2)	100.0*	0.0	0.0
	Explain in detail the duration of the experiment and the set of data that needed to be collected.	1 (1-2)	100.0*	0.0	0.0
	Show participants the orientation video.	2 (1-4)	75.0*	20.0	5.0
	Show the device to the participant in person, pointing out the different parts and what their function is to help them further understand the process.	2 (1-4)	70.0*	25.0	5.0
	Orientation to the texture and feel of the rolling device.	2 (1-5)	65.0	30.0	5.0

	Allow an upper limit of 5 unloaded practice rounds and always note in the protocol how many practice rounds were completed.	2 (1-5)	65.0	20.0	15.0
	A sensory perception (load on hand).	3 (1-5)	35.0	35.0	30.0
	Watch someone else have the measures done (if this is not in the orientation video).	3 (2-5)	30.0	25.0	45.0
Instructions for participants before the assessment	You should wear clothes that can be moved to expose your waistline. A gown or shorts might be needed.	1 (1-2)	100.0*	0.0	0.0
	You have to empty your front and back pockets including coins, keys, cellphones.	1 (1-2)	100.0*	0.0	0.0
	You should remove your glasses.	1 (1-2)	100.0*	0.0	0.0
	You should go to the restroom before testing.	1 (1-2)	100.0*	0.0	0.0
	Explain and practice breathing protocol.	1 (1-2)	100.0*	0.0	0.0
	You should disrobe/change as necessary to expose the test area sufficiently.	1 (1-2)	100.0*	0.0	0.0
	You should wear comfortable clothing.	1 (1-2)	100.0*	0.0	0.0
	Explain some circumstances where the participant might want to press the emergency stop. E.g., if they have radicular pain, and they experience pain in their leg.	1 (1-4)	95.0*	0.0	5.0
	Explain how the device works to increase participant comfort.	1 (1-3)	95.0*	0.0	5.0
	Explain how to lay down.	2 (1-4)	80.0*	5.0	15.0
	Cell phones should be allowed to stay on for emergency calls etc. but the participant should be instructed that we don't want them looking at their phones during the protocol.	2 (1-5)	65.0	15.0	20.0
Identifying the Spinous processes	Use a standardized palpation procedure based on anatomical landmarks (count up from the sacral base and down from T12/ribs) and confirm with diagnostic ultrasound.	1 (1-4)	95.0*	0.0	5.0
	Ultrasound if available.	2 (1-3)	95.0*	5.0	0.0
	Palpation in a prone position in combination with ultrasound for verification.	2 (1-4)	85.0*	10.0	5.0
	Palpation of the spinous processes.	1.5 (1-4)	85.0*	10.0	5.0
	Place hands on iliac crests, identify the L4 spinous process, place a mark on the skin, go down towards the sacrum, identify the L5 spinous process, go up towards the thoracic vertebrae, identify each spinous process.	1 (1-4)	80.0*	10.0	10.0

	Having someone with sufficient experience landmarking spinous process perform the markings.	2.5 (1-5)	50.0	25.0	25.0
	Palpation, and confirmation by a healthcare professional.	3 (1-5)	40.0	20.0	40.0
	Check by palpation done by two people.	3 (1-5)	45.0	20.0	35.0
	Identify L5 via location 1st sacral tubercle (landing point). Then L5-S1 interspinous up to L1.	3 (1-5)	45.0	30.0	25.0
	L2 spinous process is at the level of the line joining the inferior borders of the 10th ribs. The intercostal line is at the level of the L3/4 interspinous space or L3 spinous process.	3 (1-4)	35.0	40.0	25.0
	It depends on the protocol, the type of study, and the research questions being asked if accurate palpation is needed.	3.5 (1-5)	35.0	15.0	50.0
Placing the wheels over the test area	Make sure that the wheels are aligned on the skin before running each trial.	1 (1-2)	100.0*	0.0	0.0
	Make sure there is enough vertical travel in the roller to test the most posterior part of the participants' back.	1 (1-2)	100.0*	0.0	0.0
	Without changing the table height or moving the frame, move the roller wheels to the landing site by positioning the laser over the center of the "X" axis.	1 (1-2)	100.0*	0.0	0.0
	Jog wheel down onto participant and add enough cable slack (approximately 5 extra jogs down).	1 (1-2)	100.0*	0.0	0.0
	Move the roller wheels above the highest point of the test area.	1 (1-3)	95.0*	5.0	0.0
	Raise the plinth until the highest point on the participant is 3 cm from the wheels.	1 (1-3)	95.0*	5.0	0.0
	Some participants with hyper-lordosis may require more than 5 extra jogs down.	1 (1-4)	90.0*	5.0	5.0
Wheels starting position	Look at the laser from the same angle to ensure it is lined up perfectly before each trial.	1 (1-2)	100.0*	0.0	0.0
	Check the laser goes back to the reference point prior to subsequent runs.	1 (1-2)	100.0*	0.0	0.0
	Make sure the participant is not moving between the trials.	1 (1-2)	100.0*	0.0	0.0
	Mark the starting position with an "x".	1 (1-4)	95.0*	0.0	5.0
	Photos of the back should be taken.	3 (1-5)	35.0	35.0	30.0
	Measure the length of the trajectory by a tape measure.	3 (1-5)	20.0	45.0	35.0
	You should relax your back and abdominals.	1 (1-2)	100.0*	0.0	0.0

Instructions for participants during the assessment	Let us know if you wish to stop the measurements at any time or if you have any concerns (e.g., discomfort).	1 (1-2)	100.0*	0.0	0.0
	You should remain still for the duration of the test (~15 min) even when you answer a question in between the trials.	1 (1-2)	100.0*	0.0	0.0
	You will be asked to hold your breath at various times during the procedure for approximately 10 seconds each time.	1 (1-2)	100.0*	0.0	0.0
	You should wait for my instructions before you move away from the table.	1 (1-2)	100.0*	0.0	0.0
	You should keep your arm position the same for the duration of the test.	1 (1-3)	95.0*	5.0	0.0
	You should not talk during the procedure.	1 (1-4)	90.0*	5.0	5.0
	The operator should check the participant's readiness for each trial.	1 (1-4)	90.0*	5.0	5.0
	You'll be instructed when you can start breathing again.	1 (1-4)	90.0*	5.0	5.0
	You should not endure discomfort at any time especially when adding weight plates during testing.	2 (1-4)	65.0	10.0	25.0
	You should give us a sign to indicate that you have exhaled the air and ready to be tested before each trial.	2.5 (1-5)	50.0	20.0	30.0
Instructions for participants after the assessment	You should contact us if you experience any discomfort in the next few hours or days.	1 (1-2)	100.0*	0.0	0.0
	Let us know if you feel discomfort after the session or any skin irritation. These two conditions might be expected, but they will eventually disappear.	1 (1-2)	100.0*	0.0	0.0
	Wait to get up until the device is removed from above you.	1 (1-2)	100.0*	0.0	0.0
	You may experience some mild, short-term pain and discomfort in the area that has been tested.	1 (1-4)	90.0*	5.0	5.0
	You may experience some dizziness. If so, sit for a few minutes before standing up.	1 (1-4)	90.0*	5.0	5.0
	It is normal to feel slightly stiff after the measurements.	1 (1-4)	85.0*	10.0	5.0
	Slowly get up and watch your head.	1 (1-4)	85.0*	5.0	10.0
	You might feel sore in the next 48hours, this is normal but if the pain does not subside after that time or you feel worried do not hesitate to contact the principal investigator.	2 (1-5)	75.0*	15.0	10.0

	No residual pain or discomfort should remain after the measurements. Any discomfort or problems should be reported to the staff at any time.	3 (1-5)	45.0	25.0	30.0
	You should walk on a level surface (low-level exercise) for a few minutes after the test procedure.	3.5 (1-5)	15.0	35.0	50.0
	No need for specific instructions after testing. Unless there is interest in the perception of stiffness or mobility in a given study.	4 (1-5)	10.0	15.0	75.0*
A good/bad trial definition	A good trial is a trial where the wheels follow the curvature of the spine without deviating sideways, and which does not cause discomfort to the participant.	1 (1-2)	100.0*	0.0	0.0
	A good trial is one in which the participant is relaxed, does not move, and holds his/her breath out for the entire trial.	1 (1-2)	100.0*	0.0	0.0
	A bad trial is the one with irregular change in the trajectory line.	1 (1-3)	90.0*	10.0	0.0
	A good trial is consistent data collected towards a single participant.	1.5 (1-4)	85.0*	10.0	5.0
	If the wheels did not move smoothly and they are not continuously pointed forward, it is a bad trial.	2 (1-4)	70.0*	20.0	10.0
	If the displacement decreased at a higher load, it's a bad trial.	4 (2-5)	40.0	20.0	40.0
	In a good trial, the participant gets an appreciation of how the testing will feel.	4 (2-5)	10.0	30.0	60.0
	A good or bad trial would be defined based on patient reports and visual inspection.	4 (2-5)	10.0	25.0	65.0
	A good trial is when the same value is collected for all segments.	4 (2-5)	10.0	10.0	80.0*
	This is typically up to the participant whether the trial is good or bad.	5 (3-5)	0.0	10.0	90.0*
Instructions for the operator to ensure a good trial	I will monitor the wheels by enough cable slack and will align the wheels.	1 (1-2)	100.0*	0.0	0.0
	I will properly communicate with the participant what I expect from them and give them regular feedback.	1 (1-2)	100.0*	0.0	0.0
	I look for movement, breathing, and tonicity.	1 (1-2)	100.0*	0.0	0.0
	I will focus on the graphic trend.	1 (1-3)	95.0*	5.0	0.0
	I will double-check the data collected before letting the participants leave, repeat if failed.	1 (1-3)	95.0*	5.0	0.0

	I look at the graphics in the software after a few trials to make sure that the graphics look appropriate.	2 (1-5)	90.0*	5.0	5.0
	I will make sure that the graph output after each trial matches the general graph expected.	2 (1-5)	90.0*	5.0	5.0
	I'll check the values.	2 (1-5)	85.0*	5.0	10.0
	If I noticed something different with the process, I would mark it as a bad trial.	2 (1-3)	80.0*	20.0	0.0
	It is necessary that the table on which the patient is positioned has armrests to rest the arms in prone position.	2 (1-5)	75.0*	15.0	10.0
Instructions for participants for between the measurement sessions	Use the restroom.	1 (1-2)	100.0*	0.0	0.0
	Maintain your normal routine.	1 (1-2)	100.0*	0.0	0.0
	Depending on what is being investigated, researchers might need to control for exercise, food intake, hydration levels (e.g., abdominal contents, gas, delayed onset muscle soreness, etc).	1 (1-3)	95.0*	5.0	0.0
	Activities between days depending on the research question.	1 (1-3)	90.0*	10.0	0.0
	Go for a walk.	1 (1-4)	80.0*	10.0	10.0
	You must not have any treatment on the spine between sessions unless this treatment is the subject of experimentation.	2 (1-4)	80.0*	15.0	5.0
	Recommendations to be more or less active than usual could be a confounding factor to results.	2 (1-4)	75.0*	15.0	10.0
	Do not begin new physically intensive activities between measurement sessions.	1 (1-3)	75.0*	25.0	0.0
	Do not do heavy weightlifting/training in between same-day sessions.	2 (1-4)	70.0*	25.0	5.0
	If you take medication like muscle relaxants or pain killers, take the medication after the assessment.	2 (1-5)	70.0*	15.0	15.0
	No strenuous exercise should be done in between sessions.	2 (1-4)	60.0	30.0	10.0
	Come back at the same time of the day.	2 (1-4)	60.0	15.0	25.0
	Don't do any vigorous back exercises two days before the test.	2 (1-5)	55.0	25.0	20.0
	No additional care between sessions.	2.5 (1-4)	50.0	35.0	15.0
	Don't undergo any physically demanding activity involving the back.	2.5 (1-5)	50.0	15.0	35.0
	Sleep well.	3 (1-4)	40.0	45.0	15.0

	Avoid big meals in between sessions.	3 (1-4)	35.0	50.0	15.0
	Avoid swimming and scrubbing your back.	3.5 (1-5)	15.0	35.0	50.0
	Wear the same clothes for the next session.	4 (2-5)	15.0	10.0	75.0*
Instructions for the same position over multiple measurement sessions	Use a permanent marker (particularly for S1) to ensure the starting position of the measurement is the same.	1 (1-2)	100.0*	0.0	0.0
	Keep the reference points intact.	1 (1-2)	100.0*	0.0	0.0
	Have a standardized examination table with markings that could be used to align participants in a reproducible manner.	1 (1-3)	95.0*	5.0	0.0
	Take a photo with the consent of the participant.	1 (1-4)	90.0*	5.0	5.0
	Put a band-aid/ adhesive tape on top of the marked "x" spot so you don't lose it for the next visit.	1 (1-3)	85.0*	15.0	0.0
	Measure the trajectory distance.	2 (1-4)	70.0*	25.0	5.0
	Participants should feel just as comfortable as before.	2.5 (1-5)	50.0	20.0	30.0
	Since the testing plinth has a hole, the participant will always align at approximately the same distance from the cephalic end of the plinth.	3 (1-5)	45.0	40.0	15.0
	Take notes on the position of the patient (head, arms, legs).	3 (1-4)	40.0	40.0	20.0
	Tape on the table and the floor to ensure the same position of equipment and person on the table.	3 (1-5)	35.0	35.0	30.0
Software program crashes	I will stop the software and restart software.	1 (1-2)	100.0*	0.0	0.0
	I will inform the participant of the situation and will ask to lie still for the issue to be fixed.	1 (1-2)	100.0*	0.0	0.0
	I will ask the participant's permission to start over.	1 (1-2)	100.0*	0.0	0.0
	I will ask participants if they would like a rest before starting over.	1 (1-2)	100.0*	0.0	0.0
	I will re-calibrate the device.	1 (1-3)	95.0*	5.0	0.0
	I will remove all the weights.	1 (1-3)	95.0*	5.0	0.0
	Make sure the participant is safely out of the device.	1 (1-3)	95.0*	5.0	0.0
	I will remove the device from the above participant and start over.	1 (1-5)	85.0*	10.0	5.0
	My actions depend on the severity of the crash. For example, if I have to recalibrate the trajectory, I will have to recollect all trials.	1.5 (1-4)	75.0*	15.0	10.0
	I will re-do the problematic trial and resume the measurements.	2 (1-4)	70.0*	15.0	15.0

	I will re-do the measurements from 0N.	2 (1-4)	65.0	25.0	10.0
	Software program crashes are less likely to be related to the control box issue. Therefore, turning off the computer or control box will be my last resort.	2 (1-4)	65.0	25.0	10.0
	I will close the software and restart the computer.	2 (1-5)	55.0	30.0	15.0
	I will turn off the control box and restart the whole system.	2 (1-5)	55.0	15.0	30.0
	I will re-schedule the participant.	4 (1-5)	15.0	15.0	70.0*
	I will press the emergency stop button.	4 (1-5)	10.0	30.0	60.0
Participants' Safety	The safety stop button should immediately elevate the load and return the rolling arm to a position away from the patient - so that the patient can exit if needed.	1 (1-2)	100.0*	0.0	0.0
	Clear instructions to participants with expectations explained.	1 (1-2)	100.0*	0.0	0.0
	The participants should not get up before the frame is off them.	1 (1-1)	100.0*	0.0	0.0
	Make sure the device is properly operational (or locked in place) when loading weights.	1 (1-2)	100.0*	0.0	0.0
	Make sure all the 1kg weights are removed from the device before and after assessment by the VerteTrack.	1 (1-1)	100.0*	0.0	0.0
	Familiarize yourself with the location of the hardware emergency stop (E-stop) before assessment by the VerteTrack.	1 (1-2)	100.0*	0.0	0.0
	Follow the suggested pre-test protocol to make sure all "detectors" are functioning properly.	1 (1-1)	100.0*	0.0	0.0
	Procedures explained to participants for emergency stop.	1 (1-2)	100.0*	0.0	0.0
	Continuing to check in with the patient throughout the process to make sure that they are feeling okay.	1 (1-2)	100.0*	0.0	0.0
	Disinfect the wheels/bench/equipment prior to each participant.	1 (1-1)	100.0*	0.0	0.0
	Make sure to remove the weights one by one at the end of the measurement.	1 (1-4)	95.0*	0.0	5.0
	Have an easy reading format for clients with disabilities before assessment by the VerteTrack.	1 (1-4)	90.0*	5.0	5.0
	Make sure to depress the emergency stop and then disengage it to ensure it is working before assessment by the VerteTrack.	1 (1-4)	85.0*	5.0	10.0

	I will raise the plinth when not testing to make sure it will not drop if it malfunctions.	3 (1-5)	45.0	35.0	20.0
	Have a mirror to be able to see the client's face.	3 (1-5)	10.0	50.0	40.0

Table 4-3 The number of consensus statements under each category.

Category	Number of consensus statements	Number of non-consensus statements	Total number of statements
Inclusion criteria	2 (50.0%)	2 (50.0%)	4
Exclusion criteria	11 (47.8%)	12 (52.2%)	23
Pregnancy time frame limitation	0 (0.0%)	7 (100%)	7
Familiarization procedure	8 (66.7%)	4 (33.3%)	12
Instructions for participants before the assessment	10 (90.0%)	1 (9.1%)	11
Identification of spinous processes	5 (45.5%)	6 (54.5%)	11
Placing the wheels over the test area	7 (100%)	0 (0.0%)	7
Participants' starting position	4 (66.7%)	2 (33.3%)	6
Instructions for participants during the assessment	9 (81.8%)	2 (18.2%)	11
Post-test instructions	9 (81.8%)	2 (18.2%)	11
Definitions for a good or bad trial	7 (70%)	3 (30.0%)	10
Procedures to ensure a good trial	10 (100%)	0 (0.0%)	10
Instructions for between-session assessments	11 (57.9%)	8 (42.1%)	19
Instructions for reaching the same position in case of multiple assessments	6 (60%)	4 (40.0%)	10
Software program crashes	11 (68.8%)	5 (31.3%)	16
Optimizing participant safety	13 (86.7%)	2 (13.3%)	15
Total	123	60	183

4.4 Discussion

In this Delphi study, 20 panelists reached consensus on the majority of items relating to VerteTrack spinal stiffness measurements covering a wide range of domains including recruitment criteria, familiarization procedure, instructions for participants/ operators, technical issues, and safety. This

is the first time, to our knowledge, that consensus has been used to obtain a common protocol on instrumented spinal stiffness measurements.

It is important to stress that the key feature of the approach used in this study is the consensus of individuals in the field of spinal manipulative therapy and low back pain research who had experienced working with VerteTrack. Therefore, the intent was not to find "the best" protocol for measuring spinal stiffness or to present an instrument as "the only" mechanical method for measuring spinal stiffness. Our goal was to develop a standard protocol for measuring spinal stiffness using a loaded rolling wheel device that could be used as a common resource in future studies.

The surveys identified some previously known considerations when measuring stiffness including the participant's testing position, trunk muscles contraction, intra-abdominal pressure, respiratory cycle, and relocation of target spinal landmarks (Snodgrass et al., 2012; Wong & Kawchuk, 2017). This supports the quality and validity of our participants' answers as these items have been developed over years in this field and the literature. For instance, one of our participant's recommendations was to ask the patient to relax their back muscles during the assessment which is in line with an early study that showed spinal extensor muscle activities could induce changes in the mechanical responses to posteroanterior stiffness testing (Lee & Svensson, 1993). Furthermore, the surveys identified other factors not described previously in the literature including optimizing participant's safety, a definition for a good/ bad trial, procedures to ensure a good trial, placing the device over the test area, instructions for reaching the same position in case of multiple assessments, and fixing software program crashes. This emphasizes the importance of

group opinion over that of individuals for bringing new topics into focus that can be validated and studied in future works.

Interestingly, there was one specific area where no agreement was reached: the exclusion of pregnant participants from spinal stiffness measurements. One explanation for this lack of agreement is that different respondents may have different experiences in this area through diverse research designs that would, or would not, allow participants to be enrolled at different stages of pregnancy. This speculation is supported by studies to date that have employed VerteTrack. Of six studies using VerteTrack in human participants to date, three excluded pregnant participants (Fritz et al., 2020; Hadizadeh et al., 2019; Nim et al., 2020), one excluded pregnant participants in the second or third trimester of pregnancy (Brown et al., 2017) and the remaining studies did not mention pregnancy at all (Nielsen et al., 2020; Pagé & Kawchuk, 2021).

All items for which consensus was reached were consolidated into a final best practice protocol (Appendix I) for using the VerteTrack. The resulting standard protocol is expected to improve the accuracy and efficiency of spinal stiffness measurements using the VerteTrack, facilitate the training of new operators, increase consistency of these measurements in multicenter studies, and finally provide the synergy and potential for data comparison between spine studies internationally. Our final protocol provides directions for researchers and clinicians who use the VerteTrack to measure spinal stiffness. However, caution should be used if between-patient comparisons are made (for many reasons including differences in plinth rigidity as well as between-person variations). The final protocol could be useful for other technologies that assess stiffness and even manual assessment of spinal stiffness. We encourage researchers in this area to

review this protocol and consider adopting it for their own purpose. While the technical part of the protocol explaining how to operate the device may not be useful for manual assessments or devices that test participants in sitting position, however, some general information for spinal stiffness measurements has been provided and may be of benefit.

4.4.1 Strengths and limitations

The strengths of this study include the development of a consensus-based protocol based on 80% of the global population of persons with VerteTrack training and experience for Round 1 and 100% follow-up responses for Rounds 2 and 3. The relative heterogeneity in our participants may enhance the generalizability of the protocol and may have ensured that a greater spectrum of opinions was considered. The initial pilot survey improved the structure and readability of the questions before executing the full-scale project. In addition, Round 1 of our Delphi study provided the possibility of open responses and gave the participants the freedom to elaborate on the research topic which may increase the richness of the data collected. Although author bias cannot be completely eliminated from this type of research, it was minimized through implementing a Delphi consensus process using anonymous participant ratings and comments. The deidentification anonymity of participants' answers to the questions also provided more open and honest feedback and prevented response bias.

It is acknowledged that the Delphi method itself has inherent limitations including Level V in the hierarchy of evidence-based medicine and the small sample size required. Although the final protocol was developed based on Delphi participants' responses to 3 rounds of questions, it was not distributed to them for approval at the end of the study. Further, lack of interaction between

participants in the Delphi (e.g., face-to-face meetings) may deprive panelists of exchanging important information, such as clarification of reasons for disagreements.

4.5 Conclusions

Using a Delphi approach, a consensus-based protocol for measuring spinal stiffness using the VerteTrack was developed. This standard protocol was designed to i) improve the accuracy, efficiency, and safety of spinal stiffness measurements using the VerteTrack, ii) facilitate the training of new operators iii) increase consistency of these measurements in multicenter studies, and iv) provide the synergy and potential for data comparison between spine studies internationally.

Abbreviations

LBP: Low Back Pain; REDCap: Research Electronic Data Capture

Acknowledgments

The authors would like to thank all the respondents for their contribution to this consensus protocol development.

Authors' contributions

MH, GK and SF were involved in developing the design of the study. MH collected and analyzed the data in consultation with GK and SF. MH drafted the manuscript. GK and SF critically reviewed the draft manuscript. All authors read and approved the final manuscript.

Funding

Not funded.

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

This study was approved by the University of Alberta's Humans Research Ethics Board (Pro00102734) (Appendix J). All procedures performed in this study were in accordance with the ethical standards of the institutional research committee and with the declaration of Helsinki. Written informed consent was obtained from all participants in the study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

1Department of Physical Therapy, Faculty of Rehabilitation Medicine,

University of Alberta, 3-48 Corbett Hall, Edmonton, AB T6G 2G4, Canada.

2Department of Physical Therapy, Faculty of Rehabilitation Medicine,

University of Alberta, 3-44 Corbett Hall, Edmonton, AB T6G 2G4, Canada.

3Department of Chiropractic, Macquarie University, Sydney, New South Wales, Australia.

4.6 References

Akins, R. B., Tolson, H. & Cole, B. R. (2005). Stability of response characteristics of a Delphi panel: application of bootstrap data expansion. *BMC Medical Research Methodology*, *5*, 37. https://doi.org/10.1186/1471-2288-5-37

Brown, B. T., Blacke, A., Carroll, V., Graham, P. L., Kawchuk, G., Downie, A. & Swain, M. (2017). The comfort and safety of a novel rolling mechanical indentation device for the measurement of lumbar trunk stiffness in young adults. *Chiropractic and Manual Therapies*, *25*(21). https://doi.org/10.1186/s12998-017-0153-z

Colloca, C. J. (2010). The Place of Chiropractic Care in the Treatment of Low Back Pain. In M. Szpalski, R. Gunzburg, B. Rydevik, J. Le Huec & H. Mayer (Eds.), *Surgery for Low Back Pain.* (pp. 85–94). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-04547-9

Fritz, J. M., Koppenhaver, S. L., Kawchuk, G. N., Teyhen, D. S., Hebert, J. J. & Childs, J. D. (2011). Preliminary investigation of the mechanisms underlying the effects of manipulation: exploration of a multivariate model including spinal stiffness, multifidus recruitment, and clinical findings. *Spine*, *36*(21), 1772–1781. https://doi.org/10.1097/BRS.0b013e318216337d

Fritz, J. M., Sharpe, J., Greene, T., Lane, E., Hadizadeh, M., McFadden, M., Santillo, D., Farley, J., Magel, J., Thackeray, A. & Kawchuk, G. (2020). Optimization of Spinal Manipulative Therapy Protocols: A Factorial Randomized Trial Within a Multiphase Optimization Framework. *The Journal of Pain*. https://doi.org/10.1016/j.jpain.2020.11.008

Hadizadeh, M., Kawchuk, G. N. & Parent, E. (2019). Reliability of a new loaded rolling wheel system for measuring spinal stiffness in asymptomatic participants. *BMC Musculoskeletal Disorders*, *20*(1), 176. https://doi.org/10.1186/s12891-019-2543-y

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N. & Conde, J. G. (2009). Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010

Hasson, F., Keeney, S. & Mckenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, *32*(4), 1008–1015.

Iqbal, S. & Pipon-Young, L. (2009). The Delphi method. *Nursing Research*, *46*(2), 116–118.

Jun, P., Pagé, I., Vette, A. & Kawchuk, G. (2020). Potential mechanisms for lumbar spinal stiffness change following spinal manipulative therapy : a scoping review. *Chiropr Man Therap*, *28*(1), 15. https://doi.org/10.1186/s12998-020-00304-x

Keeney, S., Hasson, F. & Mckenna, H. (2006). Consulting the oracle : ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing*, *53*(2), 205–212. https://doi.org/10.1111/j.1365-2648.2006.03716.x

Keeney, S., Mckenna, H. & Hasson, F. (2011). *The Delphi technique in nursing and health research*. West Sussex, United Kingdom: Wiley-Blackwell.

Latimer, J., Lee, M., Adams, R. & Moran, C. (1996). An Investigation of the Relationship between Low Back Pain and Lumbar Posteroanterior Stiffness. *Journal of Manipulative & Physiological Therapeutics*, *19*(9), 587–591.

Lee, M. & Svensson, N. L. (1993). Effect of loading frequency on response of the spine to lumbar posteroanterior forces. *J Manipulative Physiol Ther*, *16*, 439–446.

Maher, C., Underwood, M. & Buchbinder, R. (2017). Non-specific low back pain. *The Lancet*, *389*(10070), 736–747. https://doi.org/10.1016/S0140-6736(16)30970-9

Nielsen, J., Glissmann Nim, C., O'neill, S., Boyle, E., Hartvigsen, J. & Kawchuk, G. N. (2020).

Self-reports vs. physical measures of spinal stiffness. *PeerJ*, *8*, 9598. https://doi.org/10.7717/peerj.9598

Nim, C. G., Kawchuk, G. N., Schiøttz-Christensen, B. & O'Neill, S. (2020). The effect on clinical outcomes when targeting spinal manipulation at stiffness or pain sensitivity: a randomized trial. *Scientific Reports*, *10*(1), 14615. https://doi.org/10.1038/s41598-020-71557-y

Pagé, I. & Kawchuk, G. (2021). Effects of muscle activity on lumbar spinal stiffness in asymptomatic adults: An investigation using a novel rolling device. *Musculoskeletal Science and Practice*, *52*, 102301. https://doi.org/10.1016/j.msksp.2020.102301

Snodgrass, S. J., Haskins, R. & Rivett, D. A. (2012). A structured review of spinal stiffness as a kinesiological outcome of manipulation: its measurement and utility in diagnosis, prognosis and treatment decision-making. *J Electromyogr Kinesiol*, *22*(5), 708–723. https://doi.org/10.1186/s12891-019-2543-y

Wells, C., Kolt, G. S., Marshall, P. & Bialocerkowski, A. (2014). The definition and application of pilates exercise to treat people with chronic low back pain: a Delphi survey of Australian physical therapists. *Physical Therapy*, *94*(6), 792–805. https://doi.org/10.2522/ptj.20130030

Wong, A. Y. L. & Kawchuk, G. N. (2017). The clinical value of assessing lumbar posteroanterior segmental stiffness: a narrative review of manual and instrumented methods. *PM R*, *9*(8), 816–830. https://doi.org/10.1016/j.pmrj.2016.12.001

Wong, A. Y. L., Parent, E. C., Dhillon, S. S., Prasad, N. & Kawchuk, G. N. (2015). Do participants with low back pain who respond to spinal manipulative therapy differ biomechanically from nonresponders, untreated controls or asymptomatic controls? *Spine*, *40*(17), 1329–1337. https://doi.org/10.1097/BRS.0000000000000981

Wu, A., March, L., Zheng, X., Huang, J., Wang, X., Zhao, J., Blyth, F. M., Smith, E., Buchbinder,

R. & Hoy, D. (2020). Global low back pain prevalence and years lived with disability from 1990 to 2017: estimates from the Global Burden of Disease Study 2017. *Ann Transl Med*, *8*(6), 299. https://doi.org/10.21037/atm.2020.02.175

Young, A., Swain, M. S., Kawchuk, G. N., Wong, A. Y. L. & Downie, A. S. (2020). The bench-top accuracy of the VerteTrack spinal stiffness assessment device. *Chiropr Man Ther*, *28*(1), 42. https://doi.org/10.1186/s12998-020-00331-8

Chapter 5. Predicting Who Responds to Spinal Manipulative Therapy Using a Short-Time Frame Methodology: Results From a 238-Participant Study.

A version of this chapter has been published. Hadizadeh M., Kawchuk G., Prasad N., Fritz J. "Predicting who responds to spinal manipulative therapy using a short-time frame methodology: results from a 238-participant study.", PLOS One Journal, 2020, 15 (11): 1-23.

Abstract

Background: Spinal manipulative therapy (SMT) is among the nonpharmacologic interventions that has been recommended in clinical guidelines for patients with low back pain, however, some patients appear to benefit substantially more from SMT than others. Several investigations have examined potential factors to modify patients' responses prior to SMT application. The objective of this study was to determine if the baseline prediction of SMT responders can be improved through the use of a restricted, non-pragmatic methodology, established variables of responder status, and newly developed physical measures observed to change with SMT.

Materials and methods: We conducted a secondary analysis of a prior study that provided two applications of standardized SMT over a period of 1 week. After initial exploratory analysis, principal component analysis and optimal scaling analysis were used to reduce multicollinearity

among predictors. A multiple logistic regression model was built using a forward Wald procedure to explore those baseline variables that could predict response status at 1-week reassessment.

Results: Two hundred and thirty-eight participants completed the 1-week reassessment (age 40.0± 11.8 years; 59.7% female). Response to treatment was predicted by a model containing the following 8 variables: height, gender, neck or upper back pain, pain frequency in the past 6 months, the STarT Back Tool, patients' expectations about medication and strengthening exercises, and extension status. Our model had a sensitivity of 57.4 %, specificity of 91.2%, a positive likelihood ratio of 6.5, a negative likelihood ratio of 0.5, and area under ROC curve, 0.79.

Conclusion: It is possible to predict response to treatment before application of SMT in low back pain patients. Our model may benefit both patients and clinicians by reducing the time needed to re-evaluate an initial trial of care.

5.1 Introduction

Spinal manipulative therapy (SMT) is among the nonpharmacologic interventions for low back pain (LBP) recommended as a second-line or adjunctive treatment option after exercise or cognitive behavioral therapy (Foster et al., 2018). Spinal manipulative therapy is described as a high velocity, low amplitude force applied to the vertebral column most often by chiropractors (Maitland et al., 2005). Although recommended in clinical guidelines, some patients with LBP appear to benefit substantially more from SMT than others (Childs et al., 2004). This observation has initiated several investigations that have examined potential factors to modify patients' responses prior to SMT application (Table 5-1).

Table 5-1 The previous studies examined the predictive value of baseline variables for treatment outcome in patients with low back pain receiving SMT/chiropractic treatment.

Study/ Year of publication	Study population	Baseline sample size	Type of treatment	SMT technique	Number of SMT visits	Duration of SMT program	Response assessment time	Outcome variable/ Cut off value	Possibility of prediction	Study location
Eklund A et al. 2019 (Eklund et al., 2019)	Patients with recurrent persistent LBP	593	Chiropractic treatment	Not reported	Not reported	Not reported	Fourth visit	Self-reported LBP status/ Definitely improved	Yes	Sweden
Eklund A et al. 2016 (Eklund et al., 2016)	Patients with recurrent and persistent LBP	666	Chiropractic treatment	Not reported	Not reported	Not reported	Fourth visit	Self-reported LBP status/ Definitely improved	No	Sweden
Vavrek D et al. 2015 (Vavrek et al., 2015)	Patients with chronic LBP	400	SMT/ light massage + 5 min of hot pack treatment + 5 min of very low intensity pulsed ultrasound (0.5 watts/cm2)	Pragmatic	A dose of 0, 6, 12, or 18 SMT visits	6-weeks	Shortly after completion of 6 weeks of care	≥ 50 % improvement relative to the baseline pain intensity measured by the Modified Von Korff pain scale	No	U.S.

Field J et al. 2012 (Field & Newell, 2012)	Patients with non-specific LBP	404	Not reported	Pragmatic	Not reported	Not reported	14, 30 and 90 days following the initial consultation	PGIC and BQ/ Poor outcome was defined by a PGIC response of better or much better (score of < 6), a change in total BQ score of ≤46% and a change in pain (≤ 2 points) and as derived from the pain sub-scale of the BQ	No	England
Peterson CK et al. 2012 (Peterson et al., 2012)	Patients with acute and chronic LBP	816	Chiropractic treatment	Pragmatic	Pragmatic	Pragmatic	1 week, 1 month, and 3 months after the start of treatment	The PGIC scale/ Patients responding better or much better (scores of 1 or 2) were categorized as "improved" and all other patients as "not improved."	Yes	Switzerland

Cecchi F et al. 2011 (Cecchi et al., 2011)	Patients with chronic LBP	205 (SMT group: n=69)	Booklet + advice to stay active + vertebral direct and indirect mobilization + SMT with associated soft tissue manipulation	Prescribed (Maigne, 1996)	4-6 SMT sessions (as needed) weekly sessions	4–6 once-a-week sessions. 20 minutes each session (80-120 minutes of treatment altogether)	Discharge	LBP-related functional disability assessed by RMDQ (those who decreased their RM score <2.5 were considered non-responders)	No	Not reported
Field JR et al. 2010 (Field et al., 2010)	New patients with LBP	71	Chiropractic treatment	Not reported	Not reported	Not reported	Second appointment, One month after the initial consultation	Scores > 5 on the PGIC were taken as improvement	Yes	Not reported
Leboeuf-yde C et al. 2009 (Leboeuf-Yde et al., 2009)	Patients with LBP	731	Chiropractic treatment	Not reported	Not reported	Not reported	Fourth visit, 3 months	Self-reported LBP status/ Definitely better	No	Sweden
Malmqvist S et al. 2008	New patients with LBP	984	Chiropractic treatment	Not reported	Not reported	Not reported	Second and fourth visits	The outcome (global assessment of	Yes	Finland

(Malmqvist et al., 2008)								present status at the 4th visit) was defined as positive only for those patients who reported to be definitely better at the fourth visit (or at the last visit if treatment was ended before the fourth visit).		
Langworthy JM et al. 2007 (Langworthy & Breen, 2007)	Patients with a new episode of non-specific LBP	158	Chiropractic treatment	Not reported	Not reported	Not reported	6 weeks	Deyo's Core Set/ Not reported	Yes	UK
Underwood MR et al. 2007 (Underwood et al., 2007)	Patients with LBP with a current episode duration of at least 4 weeks	1116	SMT SMT+ exercise	Prescribed (Harvey et al., 2003)	Eight sessions	12 weeks	3 months and 12 months following randomization	RMDQ score/ Not reported	No	UK

Newell D et al. 2007 (Newell & Field, 2007)	Patients with LBP	788	Chiropractic treatment	Not reported	Not reported	Not reported	4 and 12 weeks after the initial consultation	The BQ and PGIC scores/ Patients were categorised as 'better' if they chose the top two items of the scale	Yes	UK
Axén I et al. 2005 (Axén et al., 2005)	Patients with LBP	1057	Chiropractic treatment	Pragmatic	Pragmatic	Pragmatic	Fourth visit (or at the last visit if treatment was ended before the fourth visit)	Self-reported LBP status/ Definite improvement	Yes	Sweden
Axèn I et al. 2005 (Axen et al., 2005)	Patients with nonpersistent LBP	674	Chiropractic treatment	Pragmatic	Pragmatic	Pragmatic	Fourth visit	Self-reported LBP status/ Definitely improved	Yes	Sweden
Leboeuf-Yde C et al. 2005 (Leboeuf-Yde et al., 2005)	Patients with LBP	1054	Chiropractic treatment	Pragmatic	Pragmatic	Pragmatic	Fourth visit	Self-reported LBP status/ Definitely improved	Yes	Sweden

Leboeuf-yde C et al. 2005 (Leboeuf-yde & Larsen, 2005)	Patients with persistent LBP	875	Chiropractic treatment	Not reported	Not reported	Not reported	Fourth visit, 3 months and 12 months	Self-reported pain (a 0-10 box scale) and disability (the revised ODI)/ Improvement was defined as a reduction of 2 increments or more on the pain scale or as a 30% reduction in the pain score and a reduction of 20 points or more on the ODI or as a 30% reduction of the Oswestry score.	Not reported	Norway
Burton AK et al. 2004 (Burton et al., 2004)	Patients with LBP	252	Passive soft tissue stretching + passive articulation of the lumbar spine + SMT +	Not reported	Mean=6.6 sessions	Not reported	4 years	RMDQ score/ A score of 0–2 on RMDQ was considered as recovered	Yes	England

			positive encouragement + advice to stay active							
Childs JD et al. 2004 (Childs et al., 2004)	Patients with LBP	131 (SMT group: n=70)	SMT+ exercise	Prescribed (Flynn et al., 2002)	2 sessions	4 weeks	1 week	≥50% improvement in ODI	Yes	U.S.
Leboeuf-Yde C et al. 2004 (Leboeuf-Yde et al., 2004)	Patients with persistent LBP	875	Chiropractic treatment	Pragmatic	Pragmatic	Pragmatic	Fourth visit, 3 and 12 months	Maximum pain score of 1/10 and a maximum ODI score of 15/100	Yes	Norway
Axèn I et al. 2002 (Iben et al., 2002)	Patients with persistent LBP	615	Chiropractic treatment	Pragmatic	Pragmatic	Pragmatic	Fourth visit	Self-reported LBP status / Definitely improved	Yes	Sweden
Flynn T et al. 2002 (Flynn et al., 2002)	Patients with LBP	71	SMT	Prescribed	2 sessions	Treatment sessions were 2-4 days apart	Before the second and the third sessions	>50% improvement in ODI	Yes	U.S.

Skargren EI et al. 1998 (Skargren & Öberg, 1998)	Patients with low back or neck problems	323 (chiropractic group: n=179)	SMT, mobilization, traction, soft tissue treatment, instruction on individualized	Pragmatic	Mean sessions 4.9 (SD 2.0)	Mean 4.1 weeks (SD 3.3)	12 months	Mean ODI score/ Not reported	Yes	Sweden
Burton AK et al. 1995 (Burton et al., 1995)	Patients with acute and subacute LBP	252	SMT+ Exercises + general advice	Not reported	Mean sessions 6.6 (SD 5.13)	Not reported	12 months	RMDQ score/ Patients were considered recovered if they had a RMDQ score of 0-2 and not recovered if greater than 2.	Yes	England

LBP: Low Back Pain, SMT: Spinal Manipulative Therapy, ODI: Oswestry Disability Index BQ: Bournemouth Questionnaire, PGIC: Patient Global Impression of Change, RMDQ: Roland Morris Disability Questionnaire

Of these investigations, several have concluded that baseline characteristics can indeed be used to predict SMT response. A prospective study from the Nordic back pain subpopulation program examined 50 potential baseline factors in 875 LBP patients who received chiropractic care (Leboeuf-Yde et al., 2004). Their model correctly classified 99% of non-responders using 5 baseline variables: 1) sex, 2) social benefit, 3) severity of pain, 4) duration of continuous pain at first consultation, and 5) additional neck pain in the past year (Leboeuf-Yde et al., 2004). These results suggest that non-recovery from LBP in a chiropractic population is strongly related to demographic/self-report variables and weakly related to clinical variables; all five predictors were collected at the baseline without physical examination (Leboeuf-Yde et al., 2004). Interestingly though, the prediction rate for responders to chiropractic care was very low (6%). Further studies from this research group demonstrated similar results (Leboeuf-Yde et al., 2009; Leboeuf-yde & Larsen, 2005). Importantly, a subsequent validation study was performed by this group that constructed 5 predictive models on the basis of baseline information. None of the 5 models was sensitive (0-19%), whereas they were all reported highly specific (96-100%). Three factors were recognized as best at predicting non-responders by the fourth visit including no definite overall improvement by the second treatment session, the minimum total duration of LBP in the past year being 30 days, and presence of leg pain (Axén et al., 2005). Similarly, a study using a pragmatic osteopathic approach that employed SMT found two statistically significant baseline variables including depression and pain intensity as predictors of back-related disability at 4 years (Burton et al., 2004). Other studies from other groups have achieved similar results when consideration for symptom duration was given (Langworthy & Breen, 2007; Newell & Field, 2007).

Notably, a clinical prediction rule was developed to examine the characteristics of patient with LBP that may define a subgroup likely to benefit from SMT (Flynn et al., 2002). This work identified five predictive variables associated with 50% improvement in the Oswestry disability Index (ODI) within 1 week: duration of symptoms < 16 days, the fear avoidance beliefs questionnaire work subscale score < 19, at least one hip with > 35° of internal rotation range of motion, hypomobility in the lumbar spine, and no symptoms distal to the knee. According to this prospective, cohort study, patients were considered to be likely responders to manipulation when four or more of these variables were met. The probability of success with manipulation increased from 45% to 95%, when patients met this threshold. These predictive criteria was also investigated in a subsequent validation study (Childs et al., 2004). The results showed LBP patients who received manipulation and met these criteria experienced greater decreases in pain and disability after 1, 4, and 24 weeks compared to those who received manipulation but did not meet the criteria and those who met the criteria but did not receive manipulation.

On the contrary, a number of studies have had difficulty in identifying baseline characteristics of patients who respond to SMT. A secondary analysis of the large British randomized trial (UK BEAM) showed that patient baseline characteristics including age, work status, pain and disability, duration of episode, quality of life, and beliefs did not identify who was more likely to respond to manipulation or exercise with manipulation followed by exercise (combined treatment) (Underwood et al., 2007). Another retrospective analysis found that a lower baseline Roland Morris score predicted non-response to back school and individual physiotherapy but not to spinal manipulation which was provided over 4-6 weeks (Cecchi et al., 2011). In another randomized controlled trial (Vavrek et al., 2015), researchers tried to build pre- and post- treatment models to

predict responders to SMT and future pain intensity in 400 patients with chronic LBP. They reported the pre-treatment responder model in identifying SMT responders from their baseline characteristics didn't perform better than chance.

In addition, the predictive value of psychological factors in persons with LBP seeking help from chiropractors is uncertain. While an early study on the value of psychosocial variables with early identification of patients with poor prognosis showed initial psychosocial information in the form of the patient's cognitive coping strategies is highly predictive of the level of disability reported at 1 year (Burton et al., 1995), more recent studies have found little or no correlation with outcomes (Eklund et al., 2016; Field et al., 2010; Field & Newell, 2012; Langworthy & Breen, 2007; Leboeuf-Yde et al., 2009; Newell & Field, 2007).

Given the above, predicting SMT responder status at baseline may be confounded by several factors including the timeframe over which SMT applications are given, the use of additional interventions other than SMT, inclusion of treatment response variables and the choice of baseline characteristics. While many of these prior attempts at predicting SMT responder status are from pragmatic trials, application of SMT over longer time frames that reflect clinical practice may result in confounding with the natural history of the condition. Further, use of additional interventions found in clinical practice complicates interpretation and comparison between studies. Similarly, inclusion of treatment response variables voids the ability to make a baseline prediction. Finally, as our understanding of the predictive value of baseline characteristics grows, choices of which characteristics are included or excluded in the final model can cause concern.

With these issues in mind, we conducted a secondary analysis of a prior study that provided two applications of standardized SMT over a period of 1 week. The design of this prior study provides a unique opportunity to mitigate many of the potential confounders described above. Specifically, the shortened time frame of this design increases the likelihood of observing responses arising solely from SMT while decreasing the possibility of including responses associated with longer term mechanisms (e.g., natural history, contextual effects) or additional intervention. We further benefit from this design as it employs a previously validated criteria to define SMT responders; improvement in self-reported ODI occurring over 2 treatment sessions (Asher et al., 2020). Importantly, this criterion has been tied to improvements in physical measurements in responders including biomechanical, neurological and biological variables (Fritz et al., 2011; Koppenhaver et al., 2011; Wong et al., 2019) that were also collected in this study and available for use in baseline predictions. The study design also includes other new variables that have not been used previously but are increasingly thought it influence outcome (e.g. lumbar spine stiffness measures (Fritz et al., 2011; Thakral et al., 2014; Wong & Kawchuk, 2017), lumbar multifidus (LM) muscles contraction (Fritz et al., 2011; Hebert et al., 2010; Koppenhaver et al., 2011)).

Therefore, the objective of this study is to determine if the baseline prediction of SMT responders can be improved through the use of a restricted, non-pragmatic methodology, established variables of responder status, and newly developed physical measures observed to change with SMT.

5.2 Materials and methods

5.2.1 Primary protocol

In this current study, we performed a secondary analysis of data from a randomized controlled clinical trial. The original protocol for the primary study has been published previously (Fritz et al., 2018). In brief, the primary objective of the original study was to develop an optimized, multicomponent, SMT protocol using a phased, factorial design with three factors (additional SMT, multifidus muscle activation exercises, and spine mobilizing exercises). Sample size calculation was based on previous work in similar patient populations (Fritz et al., 2011). An initial sample of 280 participants was identified to provide at least 80% power to detect the minimum important differences for the patient-centered outcomes with a conservative 2-sided α=0.025 to account for co-primary outcomes. A more detailed explanation of sample size assumptions is provided in the protocol publication (Fritz et al., 2018).

Participants for the original study were individuals between 18–60 years of age with a primary complaint of LBP with or without symptoms into one or both legs, and an Oswestry disability score of at least 20%. Potential participants were excluded if they were currently receiving mind-body or exercise treatment for LBP from a healthcare provider, had "red flags" for a serious spinal condition (e.g., spinal tumor, fracture, infectious disorder, osteoporosis, or other bone demineralizing condition, etc.), showed signs consistent with nerve root compression (diminished myotomal strength, muscle stretch reflexes or sensation, positive straight leg raise), were currently pregnant, or had prior surgery to the lumbosacral spine (Appendices K and L).

After initial screening, those who provided informed consent were enrolled in the study (Appendices M and N). Each participant completed forms related to personal demographics, clinical history, and patient-reported outcomes. One of the study clinicians then performed a baseline assessment to collect various physical measurements. All participants then received two separate sessions of SMT occurring one day to one week apart. Manipulations were provided by either licensed chiropractors or physical therapists associated with the study. Following SMT, a re-assessment was conducted which collected the same baseline variables. Participants were categorized as SMT responders if their ODI score improved by 30% in 1-week reassessment.

The primary study received ethical approval from the University of Alberta (Pro00067152_Appendix O) and University of Utah (IRB_00092127_Appendix P) Institutional Review Boards. All the patients' data were fully anonymized. Permission to use anonymized data for the present study was obtained by the responsible authority, Julie M Fritz.

5.2.2 Demographic and history measures

Basic demographic information including age, gender, race, ethnicity, weight, height, marital status, employment status, highest education level, and clinical history (e.g., duration of symptoms, comorbid health conditions, prior history of LBP) were collected.

5.2.3 Patient reported outcome measures

Baseline assessment also included the ODI (Appendix Q) and Numeric Pain Rating Scale (NPRS) which were used as participant self-report measures of function and pain respectively (Childs et

al., 2005; Fritz & Irrgang, 2001). The Fear-Avoidance Beliefs Questionnaire (FABQ) was also collected to measure patient beliefs about how physical activity and work may affect their LBP and perceived risk for re-injury (Appendix R) (Waddell et al., 1993). In addition, short forms from the University of Washington concerns about pain (UWCAP) and pain-related self-efficacy (UWPRSE) item banks were collected to measure the extent to which people catastrophize in response to pain and their degree of confidence in the ability to function with pain respectively (Appendix S). We also assessed the participant's risk of persistent disabling pain as low, medium, or high risk using the STarT Back Tool (SBT) (Hill et al., 2008). Patients were asked about their expectations of LBP outcomes specifically related to medications, surgery, rest, X-ray, MRI, modalities, traction, manipulation, massage, strengthening, aerobic, and range of motion exercises.

5.2.4 Physical examination measures

Physical examination measures included assessment of spinal (flexion, extension, left and right side-bending) (Waddell et al., 1992) and hip range of motion (left and right internal rotation), lumbar segmental testing for mobility with manually applied posterior-anterior force (Maher et al., 1998), pain on palpation, straight leg raise (SLR) (Magee, 2007), Aberrant movements during lumbar range of motion (Fritz et al., 2006), multifidus lift test at two levels (L4-L5 and L5-S1) and a prone instability test (Fritz et al., 2006; Magee, 2007).

5.2.5 Instrumented measures

Both LM muscle activation and lumbar spine stiffness were evaluated at the baseline. Multifidus activation was measured with brightness-mode ultrasound images using a Sonosite MicroMaxx (Sonosite Inc. Bothell, WA, USA) and a 60-mm, 2–5 MHz curvilinear array transducer based on

a previously validated protocol (Kiesel et al., 2007). Participants were positioned prone with their head neutral and a pillow under their abdomen to flatten the lordosis. Images were obtained at two vertebral levels (L4-L5 and L5-S1) in the parasagittal plane during rest (static) and submaximal contraction (dynamic) in response to the participant lifting a small weight with the contralateral hand. The weight was selected according to the participant's mass (<150 lb: 1.5 lb; 150-200 lb: 2 lb; and >200 lb: 3 lb). Three images were acquired in each state (relaxed and contracted) for each side and at two levels (L5/S1, L4/5), one side at a time. Images were stored and analyzed offline using ImageJ V1.38t software (National Institutes of Health, Bethesda, MD). Offline measures of LM thickness were obtained from determining the distance between the posterior-most aspect of the facet joint inferiorly and the plane between the multifidus and thoracolumbar fascia superior for both the resting and contracted states. Multifidus muscle activation was calculated as: (Thickness $_{contracted}$ – Thickness $_{relaxed}$) / Thickness $_{relaxed}$) (Kiesel et al., 2007). The average of three measures was used for the analysis, for the total of 8 variables.

Lumbar spinal stiffness was assessed with the VerteTrack™ (VibeDx Corporation, Canada) which uses a rolling wheel system to apply vertical loads over the spine of a prone participant. The VerteTrack houses multiple sensors to provide continuous, real-time quantification of spinal deformation in response to a defined load. The resulting force displacement curves were used to calculate stiffness at each lumbar segment in N/mm. Terminal Stiffness was calculated as the ratio of the maximum applied force to the resultant displacement at each lumbar level (Fritz et al., 2011). Global stiffness was determined from the slope of force-displacement curve between 5 N and 60 N, representing the stiffness of underlying tissues throughout each trial (Fritz et al., 2011). One measure per lumbar segment corresponding to general stiffness, terminal stiffness, last load, and

displacement were retained for analysis, for a total of 20 variables. The within- and between-session reliability and accuracy for spinal stiffness measures taken with this device has been evaluated previously (Hadizadeh et al., 2019; Young et al., 2020).

5.2.6 Spinal manipulative therapy

All SMT sessions began with a brief assessment by the clinician to identify possible SMT contraindications. The preferred SMT technique has been described previously (Childs et al., 2004). This procedure is performed with the participant supine. The clinician stands opposite the side to be manipulated and side-bended the participant. The side to be manipulated was the side identified as more painful on the basis of participant's report. If the participant couldn't identify a more painful side the clinician selected a side. The participant crossed their arms in front of the chest while the clinician rotated him/her and delivered a high-velocity, low-amplitude (HVLA) thrust to the anterior superior iliac spine in a posterior/inferior direction.

If this technique was not possible due to participant preference or comfort, a side-posture HVLA was performed. The participant laid on their uninvolved side with their superior leg bent to 90˚ and the clinician places their pisiform on to their posterior superior iliac spine and delivers a high velocity low amplitude (HLVA) thrust. Previous study found no difference in outcome between this SMT procedure and a side-posture HVLA technique (Cleland et al., 2009) while both techniques have been found to be well-tolerated (Cleland et al., 2009).

Spinal manipulative therapy was considered complete if a cavitation (i.e., a "pop") occurred following SMT application. If cavitation was not achieved, the participant was repositioned and SMT performed again. If no cavitation occurred on this second attempt, the clinician performed SMT on the opposite side. A maximum of 2 attempts per side was permitted. If no cavitation was noted after the fourth attempt, SMT was complete. The number of SMT attempts and the technique used were recorded by the clinician.

5.2.7 Statistical analysis

All measures collected at baseline were used at the beginning of this analysis. Continuous data was summarized by means, medians and standard deviation. Categorical data was summarized by frequencies and percentages.

We have summarized the statistical methods used for data analysis in Figure 5-1. An initial exploratory analysis demonstrated that the collected variables at the baseline were associated with the relative changes in ODI. However, a high correlation was found between most of the ultrasound values, stiffness measures, and lumbar mobility testing results in bivariate correlation analysis ($R \geq \pm 0.7$), therefore a principal component analysis using varimax rotation with Kaiser normalization was conducted to address this multicollinearity and reduce the number of variables input into the subsequent multiple regression model (Underwood et al., 2007). An optimal scaling analysis was also performed to address the problem of too few observations for some of the categorical variables. Optimal scaling is a general approach to treat multivariate data through the optimal transformation of qualitative scales to quantitative values. Using this approach, both nominal and ordinal variables can be optimally transformed into numerical values to reduce multicollinearity

among predictors and maximize the homogeneity or internal consistency among variables. As a result nonlinear relationships between transformed variables can be modeled (McCormick et al., 2017; Meulman, 2000). Finally, a multiple logistic regression model was built using a forward Wald procedure to explore those baseline variables that could predict overall outcome (response status) at 1-week reassessment (Vavrek et al., 2015). Analyses were conducted using IBM SPSS version 26.0 (Armonk, New York, USA). An alpha value of 0.05 was used for all analysis. In addition, sensitivity/specificity, positive/negative predictive values, positive/ negative likelihood ratios (Fritz & Wainner, 2001), and the area under the receiver operating characteristic (ROC) curve were estimated for the final model.

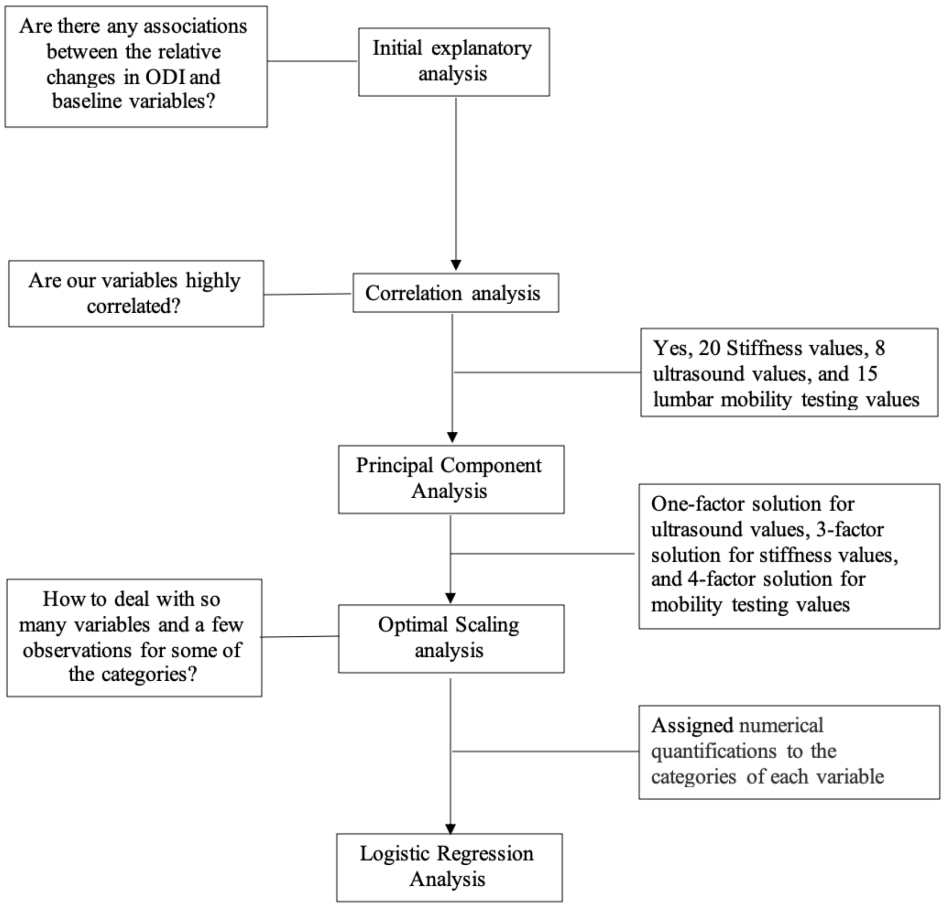


Figure 5-1 Statistical analysis for prediction of response to spinal manipulative therapy

5.3 Results

Two hundred and thirty-eight participants completed the 1-week reassessment (age 40.0± 11.8 years; 59.7% female). Table 5-2 to Table 5-6 present the results of the history and demographic, patient-reported outcome measures, patients' expectations, physical examination and instrumented measures at the baseline, respectively.

Table 5-2 History and demographic variables assessed at baseline.

Characteristics	**All Participants (n=238)**	**Responders (n=68)**	**Non-responders (n=170)**
Age (y)	40.0± 11.8	40.4± 10.8	39.8± 12.2
Sex (% female)	59.7	57.4	60.6
Race (%)			
American Indian or Alaskan Native	1.7	0.0	2.4
Asian	10.5	11.8	10.0
Black or African American	2.9	4.4	2.4
White or Caucasian	73.1	64.7	76.5
Other	6.3	13.2	3.5
> one race	5.5	5.9	5.3
Ethnicity (%)			
Hispanic or Latino	8.4	13.2	6.5
Not Hispanic or Latino	91.6	86.8	93.5
Marital status (%)			
Single, widowed, or divorced	36.6	30.9	38.8
Married	51.7	60.3	48.2
Live with significant other	11.8	8.8	12.9
Height (cm)	170.9± 10.4	168.9± 10.4	171.7± 10.4
Body mass index (kg/m2)	28.4± 7.0	27.5± 6.7	28.8± 7.1
Education level (%)			
Did not complete high school	2.1	1.5	2.4
Completed high school	34.9	23.5	39.4
Completed college degree	63.0	75.0	58.2
Current work status (%)			
Not employed outside the home	15.5	19.1	14.1
Employed part-time	17.2	16.2	17.6
Employed full-time	59.2	61.8	58.2
Not employed for low back condition	5.9	1.5	7.6
Retired	2.1	1.5	2.4
Workers' compensation (% yes)	3.4	4.4	2.9

Prior history of LBP (% yes)	61.8	60.3	62.4
Pain Duration	4000.0± 4149.0	3247.0± 3534.8	4301.1± 4343.7
Duration of current symptoms (d)	1116.5± 2312.4	1203.0± 2587.1	1082.0± 2200.3
LBP Frequency in the past 6 months (%)			
Every day or nearly every day	65.5	57.4	68.8
At least half the days	16.4	7.4	20.0
Less than half the days	18.1	35.3	11.2
Distal-most extent of symptoms (%)			
Low back only	41.2	38.2	42.4
Buttock(s)	37.4	48.5	32.9
Thigh(s) - above the knee	15.5	11.8	17.1
Below the knee(s)	5.9	1.5	7.6
Current medications regular usage for back pain (% yes)			
Acetaminophen	15.9	11.8	17.7
Non-Steroidal Anti-Inflammatories	26.5	19.2	20.4
Steroids	0.0	0.0	0.0
Opioid	6.7	4.4	7.7
Other	10.5	4.4	13.0
Comorbid health conditions (% yes)			
Diabetes	5.0	4.4	5.3
High Blood Pressure	8.0	7.4	8.2
Cancer	0.0	0.0	0.0
Depression	21.4	8.8	26.5
Anxiety	23.9	14.7	27.6
Other mental health condition	6.7	1.5	8.8
Rheumatoid arthritis	2.1	0.0	2.9
Neck or upper back pain	25.6	13.2	30.6
Substance or alcohol abuse	0.0	0.0	0.0
Cigarette Smoking history (%)			
Non-smoker	64.7	72.1	61.8
Ex-smoker	21.0	16.2	22.9
Current smoker	14.3	11.8	15.3
Previous tests (% yes)			

X-rays	57.1	45.6	61.8
MRI	27.3	22.1	29.4
CT scan	8.4	4.4	10.0
Other imaging	2.5	2.9	2.4
None	38.2	50.0	33.5
Treatment Used for LBP Episode (%yes)			
Chiropractic	46.6	35.3	51.2
Physical Therapy	40.8	30.9	44.7
Steroid Injections	13.4	8.8	15.3
Corset/Brace	8.4	10.3	7.6
Opioid Medication	19.7	14.7	21.8
Massage Therapy	37.8	26.5	42.4
Cognitive Behavioral	3.8	2.9	4.1
Therapy/Counseling	34	36.8	32.9
Other	20.6	29.4	17.1

NOTE. Values are mean ± SD unless otherwise indicated.

Table 5-3 Patient-reported outcome measures at baseline.

Characteristics	**All Participants (n=238)**	**Responders (n=68)**	**Non-responders (n=170)**
Numeric pain rating scale (0-10)	4.6± 1.6	4.2± 1.7	4.8± 1.6
Oswestry disability index (0-100)	34.1± 11.8	34.0± 12.8	34.1± 11.4
Psychosocial covariate measures			
Short form UWCAP	49.2± 8.9	49.0± 8.1	50.5± 8.9
Short form UWPRSE	51.6± 8.2	53.3±7.5	50.9± 8.3
FABQ score (0-96)			
Work subscale (0-42)	15.6± 10.0	13.9± 9.1	16.3± 10.2
Physical activity subscale (0-24)	14.5± 4.9	14.0± 4.8	14.7± 4.9
SBT total score	4.3± 1.9	3.8± 1.8	4.6± 1.9
SBT psychological distress score	2.3± 1.4	2.03± 1.2	2.4± 1.4
SBT categorization (%)			
Low risk	33.2	44.1	28.8
Medium risk	46.2	45.6	46.5
High risk	20.6	10.3	24.7

NOTE. Values are mean ± SD unless otherwise indicated.

Numeric pain rating scale reports the average of the worst, best, and current scores for pain over the last 24 hours using a self-reported 0-10 numerical pain rating scale ranging from '0' no pain, and '10' worst imaginable pain (Childs et al., 2005). Function was evaluated using Oswestry Disability Index on a 0-100 scale, with lower numbers indicating better function (Fritz & Irrgang, 2001). Fear-avoidance beliefs about physical activity and work were assessed using the Fear Avoidance Beliefs Questionnaire (FABQ) (Waddell et al., 1993). The short form of the University of Washington concerns about pain (UWCAP) is a measure of pain catastrophizing including 8-items, with each item rated on a 5-point scale: 1 (Never) to 5 (always). The higher the score, the more catastrophizing thoughts are present. The short form of the University of Washington pain-related self-efficacy (UWPRSE) was used to assess one's confidence in performing particular activities while in pain. It is a 9-item scale, with each item rated on a 5-point scale: 0 (Not at all) to 5 (very much). Higher scores represent higher confidence to function with pain. The short forms of the UWCAP and the UWPRSE items were scored by converting the total raw score into an item response theory-based T-score for with a mean of 50 and a standard deviation of 10. The mean score of 50 represents a mean of a large sample of people with chronic pain. The STarT Back Tool (SBT) is a 9-item questionnaire including physical and psychosocial statements that are used to categorize patients into low, medium, or high-risk groups for persistent LBP-related disability (Hill et al., 2008).

Table 5-4 Patient expectations about different interventions at baseline.

Patients' expectations (%)	**All Participants (n=238)**	**Responders (n=68)**	**Non-responders (n=170)**
Medications			
Completely disagree	10.9	13.2	10.0
Somewhat disagree	18.5	22.1	17.1
Neutral	24.4	33.8	20.6
Somewhat agree	42.0	26.5	48.2
Completely agree	4.2	4.4	4.1
Surgery			
Completely disagree	36.6	50.0	31.2
Somewhat disagree	18.9	14.7	20.6
Neutral	33.6	29.4	35.3
Somewhat agree	9.2	5.9	10.6
Completely agree	1.7	0.0	2.4

Rest			
Completely disagree	12.6	8.8	14.1
Somewhat disagree	11.8	10.3	12.4
Neutral	18.1	20.6	17.1
Somewhat agree	44.5	50.0	42.4
Completely agree	13.0	10.3	14.1
X-ray			
Completely disagree	16.4	16.2	16.5
Somewhat disagree	15.5	19.1	14.1
Neutral	40.3	35.3	42.4
Somewhat agree	19.3	20.6	18.8
Completely agree	8.4	8.8	8.2
MRI			
Completely disagree	11.8	13.2	11.2
Somewhat disagree	11.8	16.2	10.0
Neutral	37.8	33.8	39.4
Somewhat agree	29.0	27.9	29.4
Completely agree	9.7	8.8	10.0
Modalities			
Completely disagree	1.7	0.0	2.4
Somewhat disagree	2.9	2.9	2.9
Neutral	8.8	11.8	7.6
Somewhat agree	59.2	57.4	60.0
Completely agree	27.3	27.9	27.1
Traction			
Completely disagree	6.7	7.4	6.5
Somewhat disagree	3.8	4.4	3.5
Neutral	42.9	45.6	41.8
Somewhat agree	37.0	32.4	38.8
Completely agree	9.7	10.3	9.4
Manipulation			
Completely disagree	3.4	2.9	3.5
Somewhat disagree	4.2	5.9	3.5
Neutral	18.1	19.1	17.6
Somewhat agree	55.0	48.5	57.6
Completely agree	19.3	23.5	17.6
Massage			

Completely disagree	2.5	4.4	1.8
Somewhat disagree	3.8	0.0	5.3
Neutral	8.4	13.2	6.5
Somewhat agree	51.7	44.1	54.7
Completely agree	33.6	38.2	31.8
Strengthening exercises			
Completely disagree	0.8	0.0	1.2
Somewhat disagree	2.1	1.5	2.4
Neutral	6.3	4.4	7.1
Somewhat agree	39.1	30.9	42.4
Completely agree	51.7	63.2	47.1
Aerobic exercises			
Completely disagree	5.0	1.5	6.5
Somewhat disagree	10.9	11.8	10.6
Neutral	22.7	32.4	18.8
Somewhat agree	41.6	30.9	45.9
Completely agree	19.7	23.5	18.2
Range of motion exercises			
Completely disagree	0.8	1.5	0.6
Somewhat disagree	1.7	1.5	1.8
Neutral	7.6	5.9	8.2
Somewhat agree	42.0	38.2	43.5
Completely agree	47.9	52.9	45.9

Table 5-5 Physical examination variables assessed at baseline.

Variable	All subjects (n=238)	Responders (n=68)	Non-responders (n=170)
Range of Motion			
Right side-bending (°)	25.4± 8.8	25.6± 8.7	25.3± 8.9
Left side-bending (°)	25.8± 8.9	26.8± 8.2	25.4± 9.1
Total flexion (°)	91.2± 24.3	95.2± 21.5	89.6± 25.3
Total extension (°)	24.5± 10.7	22.9± 10.5	25.1± 10.7
Right hip internal rotation (°)	31.0± 11.7	30.0± 11.9	31.4± 11.7
Left hip internal rotation (°)	31.0± 11.6	30.8± 11.8	31.1± 11.5
Right side-bending status (%)			
Centralized	16.8	10.3	19.4
Status Quo	76.5	86.8	72.4
Peripheralized	6.7	2.9	8.2
Left side-bending status (%)			
Centralized	12.6	5.9	15.3
Status Quo	80.7	89.7	77.1
Peripheralized	6.7	4.4	7.6
Total flexion status (%)			
Centralized	13.4	7.4	15.9
Status Quo	78.6	88.2	74.7
Peripheralized	8.0	4.4	9.4
Total extension status (%)			
Centralized	19.3	10.3	22.9

Status Quo	75.2			82.4			72.4		
Peripheralized	5.5			7.4			4.7		
Additional Tests									
Right straight leg raise test (°)	73.5± 14.5			72.9± 12.2			73.8± 15.4		
Left straight leg raise test (°)	72.3± 16.1			73.3± 13.1			71.9± 17.1		
Aberrant movements during ROM (% Positive)	37.4			45.6			34.1		
Multifidus lift test L4/L5 (% Abnormal)	35.3			36.8			34.7		
Multifidus lift test L5/S1 (% Abnormal)	39.5			45.6			37.1		
Prone instability test (% Positive)	21.4			26.5			19.4		
Manual Mobility Assessment (%)	**Hypomobile**	**Norm**	**Hypermobile**	**Hypomobile**	**Norm**	**Hypermobile**	**Hypomobile**	**Norm**	**Hypermobile**
L1 mobility	32.8	63.4	3.8	29.4	64.7	5.9	34.1	62.9	2.9
L2 mobility	34.0	61.8	4.2	32.4	61.8	5.9	34.7	61.8	3.5
L3 mobility	46.2	49.6	4.2	44.1	50.0	5.9	47.1	49.4	3.5
L4 mobility	58.8	36.1	5.0	60.3	36.8	2.9	58.2	35.9	5.9
L5 mobility	63.4	33.2	3.4	58.8	39.7	1.5	65.3	30.6	4.1
Pain on palpation (% yes)									
L1 pain	32.4			27.9			34.1		
L2 pain	43.7			39.7			45.3		
L3 pain	56.3			57.4			55.9		
L4 pain	67.6			67.6			67.6		
L5 pain	67.6			57.4			71.8		

NOTE. Values are mean ± SD unless otherwise indicated.

Table 5-6 Instrumented measures at baseline.

Characteristics	All Participants		Responders		Non-responders	
Multifidus Activation						
Right L4_L5	3.8± 1.1		3.8± 1.0		3.8± 1.1	
Left L4_L5	3.8± 1.1		3.8± 1.1		3.8± 1.1	
Right L5_S1	3.6± 1.1		3.6± 1.1		3.6± 1.2	
Left L5_S1	3.7± 1.2		3.7± 1.2		3.7± 1.1	
Multifidus Rest						
Right L4_L5	3.4± 1.1		3.4± 1.1		3.4± 1.1	
Left L4_L5	3.5± 1.1		3.5± 1.2		3.5± 1.1	
Right L5_S1	3.3± 1.2		3.4± 1.2		3.3± 1.2	
Left L5_S1	3.4± 1.2		3.5± 1.3		3.4± 1.2	
Spinal Stiffness (N/mm)	**Global**	**Terminal**	**Global**	**Terminal**	**Global**	**Terminal**
L1	4.5± 1.0	5.8± 1.1	4.6± 1.0	5.9± 1.1	4.5± 1.0	5.8± 1.2
L2	4.4± 0.9	5.7± 1.1	4.5± 1.0	5.8± 1.1	4.4± 0.9	5.7± 1.1
L3	4.4± 0.9	5.7± 1.1	4.6± 0.9	5.9± 1.1	4.4± 0.8	5.6± 1.0
L4	4.5± 0.9	5.8± 1.2	4.7± 1.0	6.1± 1.2	4.5± 0.9	5.7± 1.1
L5	4.7± 1.1	6.0± 1.3	4.9± 1.1	6.3± 1.4	4.6± 1.0	5.9± 1.3

NOTE. Values are mean ± SD.

Principal component analysis identified a three-factor solution for the stiffness values, one-factor solution for ultrasound values, and four-factor solution for the mobility testing results. Together these factors explained 89.1%, 90.1%, and 78.3% of the variance in the stiffness, ultrasound, and lumbar mobility testing data respectively. Lumbar spine stiffness values, LM activation values, and mobility testing results were then converted into principal component scores to construct our model.

Logistic regression analysis resulted in a model with eight baseline variables (Table 5-7). The 8 variables in this model represent a number of different domains including participant demographics (height and gender), history (neck or upper back pain and pain frequency in the past

6 months), participant self-reported measures (SBT, patients' expectations about medication and strengthening exercises) and physical examination (extension status). Two variables were removed: One variable (depression) for not being statistically significant (P-value> 0.05) and another one (current pain duration) for having a regression coefficient of 0 and odds ratio (OR) equals to 1 showing there was no difference between responders and non-responders in the duration of their current pain.

Table 5-7 Logistic regression analysis of 238 participants with low back pain for relative changes in Oswestry disability index following spinal manipulative therapy resulting in an 8-variable model.

Predictor	**ß**	**Std. Error**	**Wald**	**P-Value**	**Odds ratio ($e^{ß}$)**	**95% Confidence Interval**		**Interpretation**
						Lower limit	**Upper limit**	
Height	-0.29	0.07	16.13	0.00	0.75	0.65	0.86	Shorter, more improvement
Gender	-0.87	0.28	11.41	0.00	0.42	0.24	0.73	Male, more improvement
Current pain duration	0.00	0.00	6.35	0.01	1.00	1.00	1.00	No changes
Depression	-0.39	0.22	3.32	0.07	0.68	0.44	1.03	Not significant
Neck or upper back pain	-0.63	0.21	9.25	0.00	0.53	0.35	0.80	No neck or upper back pain, more improvement
Pain frequency in the past 6 months	0.81	0.18	20.23	0.00	2.25	1.58	3.20	More pain frequency, more improvement
Patient's expectation on medication	-0.72	0.20	13.23	0.00	0.49	0.33	0.72	Lower expectation, more improvement
Patient's expectation on strengthening exercises	0.90	0.35	6.60	0.01	2.47	1.24	4.93	Higher expectation, more improvement
The STarT Back Tool	-0.31	0.10	8.80	0.00	0.74	0.60	0.90	Lower score, more improvement
Extension status	0.39	0.18	4.77	0.03	1.48	1.04	2.11	Peripheralized pain with extension, more improvement

As seen in Table 5-7, the effect of gender is significant but negative, indicating that females were 0.42 times less likely to respond to SMT than males. Higher expectations about strengthening (OR= 2.47) was associated with an increased likelihood of responding to SMT but higher expectation about medication (OR= 0.49) was associated with a reduction in the likelihood of responding to SMT. Participants with peripheralized pain during extension and those with more frequent pain in the past six month were 1.48 and 2.25 times more likely to be SMT responders, respectively. The ß coefficient for height, neck or upper back pain, and SBT score were also significant and negative indicating that increasing affluence is associated with decreased odds of responding to treatment.

Table 5-8 presents the degree to which predicted probabilities agree with actual outcomes in a classification table. The overall correct prediction, 81.5% shows an improvement over the chance level which is 50%. Our model had a sensitivity of 57.4%, specificity of 91.2%, a positive likelihood ratio of 6.5, a negative likelihood ratio of 0.5, and area under ROC curve, 0.79.

Table 5-8 The observed and the predicted frequencies for responders and non-responders to spinal manipulative therapy by logistic regression for the final model with the cut off of 0.50.

	Observed		
Predicted	**Non-responder**	**Responder**	**Overall% Correct**
Responder	39	15	
Non-responder	29	155	
% correct	57.4	91.2	81.5

Note. Sensitivity = 39/ (39+ 29) %= 57.4%. Specificity =155/ (155+15) %= 91.2%. Positive predictive value = 39/ (39+15) %= 72.2%. Negative predictive value = 155/ (155+29) %= 84.2%.

5.4 Discussion

Identification of SMT responders and non-responders prior to application of the SMT has received increasing attention in the conservative treatment of patients with LBP; however, the evidence for the effectiveness of this approach is mixed. To determine if the baseline prediction of SMT responders can be improved through the use of a restricted, non-pragmatic methodology, established definitions of responder status, and newly developed physical measures observed to change with SMT, we investigated the predictive values of 20 history and demographic variables, 6 patient-reported outcome measures, 22 physical measures, and 28 instrumented measures as unique domains and in combination. Our results suggest that it is possible to predict SMT response in a specific group of patients with 91.2% accuracy in non-responder and 57.4% in responder after only two applications of standardized SMT over a one-week period. To our knowledge, this is the first investigation to achieve prediction results of this magnitude for responder group although the model has yet to be validated.

Prior studies that have generated successful predictions of SMT response have tended to arise from pragmatic designs. In contrast, prior studies that have chosen to provide SMT alone or with minimal additional interventions have not achieved successful predictions. While it is possible that the prior success of pragmatic studies in this regard is because a pragmatic design more closely mimics clinical practice, our results do not support that idea. Specifically, our methodology applied fewer SMTs over a shorter time frame using a pre-defined technique for SMT application. Therefore, one explanation for our non-congruent results is that our hypothesis is tenable; that is, predicting SMT response is best assessed in a short-time frame and in isolation of other interventions.

In addition, the magnitude of our SMT responder prediction was substantially greater when compared to prior studies that have not exceeded 19% to date. In the clinical prediction rule developed by Flynn et al, SMT response was predicted with 100% in non-responders and 19% in responders (Flynn et al., 2002). Although this previous model consisted of fewer variables (i.e., 5) that is presumably easier to manage, the prediction performance for responders was lower. While at first glance it may appear unwieldy to use an 8-variable model including a 9-item questionnaire in a future clinical situation, 7 of the 8 variables can be collected in advance of the examination. The remaining one variable can be collected by clinicians with relative ease and expediency (extension status). In addition, one fourth of the model presented in the study is about patients' expectations on treatment. Although previous studies showed illness beliefs and beliefs about rehabilitation make a significant contribution to the prediction of different rehabilitation outcome indicators, the reason for this association remains unexplained (Glattacker et al., 2013; Hayden et al., 2019; Iles et al., 2009; Kongsted et al., 2014; Metcalfe & Moffett, 2005; Myers et al., 2008). However, it would be worthwhile to address the power of treatment expectations in comparison to other psychosocial factors in this group of patients. Importantly, none of the clinical measures included in our final model involved newly described physical measures involving special equipment and training (ultrasonic evaluation of muscle contraction, evaluation of spinal stiffness evaluation with a mechanical device).

The strengths of our study include a multi-site design which would tend to mitigate the possibility of our results arising from a specific population. Although most previous studies used other measures as response criteria, we defined our response value as 30% improvement on the ODI

which is an accepted threshold of change based on minimal clinically important difference scores for this questionnaire (Gatchel & Mayer, 2010; Ostelo et al., 2008). Given this and considering the moderate sensitivity and high specificity of our prediction results, we propose that a future validation study of this model is warranted. If found to be valid, these 8 variable models could provide clinicians with the opportunity to construct a more focused intervention plan after only 1 week of care. This would benefit both patients and clinicians by reducing more traditional re-evaluation periods of an initial trial of care that may extend into multiple weeks with many more treatment sessions.

As with all experiments, our study had limitations. First, our sample was heterogeneous in terms of pain duration. Although most participants in this study could be classified as having chronic LBP, our inclusion criteria were not limited to chronicity. Since the original primary study was designed to assess therapeutic effects in a wide range of participants, it did not restrict enrollment to a specific duration of low back pain. Therefore, the usability of the proposed model cannot be easily extrapolated to populations that may be highly homogeneous in pain duration. Second, we did not have a control group, thus these outcome data cannot be regarded as a clinical prediction rule, however, it can inform the professions of what might be important in patients' clinical assessment.

5.5 Conclusion

The 8 variable model presented here was able to predict SMT response with a sensitivity of 57.4% a specificity of 91.2%, and an overall classification accuracy of 81.5%. Given these results, and that 7 model variables can be collected prior to clinician engagement, future validation of the

model is warranted. Should the model be valid, it may benefit both patients and clinicians by reducing the time needed to re-evaluate an initial trial of care.

Acknowledgments

The authors would like to thank Dr. Randy Vollrath, Dr. Shannon Wandler, Dr. Moe Gebara, Dr. Lacina Barsalou, Lynne Wong, and Erica Marr for their assistance during the data collection phase.

5.6 References

Asher, A. M., Oleisky, E. R., Pennings, J. S., Khan, I., Sivaganesan, A., Devin, C. J., Bydon, M., Asher, A. L. & Archer, K. R. (2020). Measuring clinically relevant improvement after lumbar spine surgery: is it time for something new? *Spine Journal*, *20*. https://doi.org/10.1016/j.spinee.2020.01.010

Axén, I., Jones, J. J., Rosenbaum, A., Lövgren, P. W., Halasz, L., Larsen, K. & Leboeuf-Yde, C. (2005). The nordic back pain subpopulation program: Validation and improvement of a predictive model for treatment outcome in patients with low back pain receiving chiropractic treatment. *Journal of Manipulative and Physiological Therapeutics*, *28*(6), 381–385. https://doi.org/10.1016/j.jmpt.2005.06.008

Axen, I., Rosenbaum, A., Robech, R., Larsen, K. & Leboeuf-Yde, C. (2005). The Nordic back pain subpopulation program: Can patient reactions to the first chiropractic treatment predict early favorable treatment outcome in nonpersistent low back pain? *Journal of Manipulative and Physiological Therapeutics*, *28*(3), 153–158. https://doi.org/10.1016/j.jmpt.2005.02.007

Burton, A. K., McClune, T. D., Clarke, R. D. & Main, C. J. (2004). Long-term follow-up of patients with low back pain attending for manipulative care: Outcomes and predictors. *Manual Therapy*, *9*(1), 30–35. https://doi.org/10.1016/S1356-689X(03)00052-3

Burton, A. K., Tillotson, K. M., Main, C. J. & Hollis, S. (1995). Psychosocial predictors of outcome in acute and subchronic low back trouble. *Spine*, *20*(6), 722–728. https://doi.org/10.1097/00007632-199503150-00014

Cecchi, F., Negrini, S., Pasquini, G., Paperini, A., Conti, A., Chiti, M., Zaina, F., Macchi, C. & Molino-Lova, R. (2012). Predictors of functional outcome in patients with chronic low back pain undergoing back school, individual physiotherapy or spinal manipulation. *European*

Journal of Physical and Rehabilitation Medicine, *48*(3), 371–378.

Childs, J. D., Fritz, J. M., Flynn, T. W., Irrgang, J. J., Johnson, K. K., Majkowski, G. R. & Delitto, A. (2004). A Clinical Prediction Rule To Identify Patients with Low Back Pain Most Likely To Benefit from Spinal Manipulation: A Validation Study. *Annals of Internal Medicine*, *141*(12), 920. https://doi.org/10.7326/0003-4819-141-12-200412210-00008

Childs, J. D., Piva, S. R. & Fritz, J. M. (2005). Responsiveness of the Numeric Pain Rating Scale in Patients with Low Back Pain. *Spine*, *30*(1), 1331–1334. https://doi.org/10.1097/01.brs.0000164099.92112.29

Cleland, J. A., Fritz, J. M., Kulig, K., Davenport, T. E., Eberhart, S., Magel, J. & Childs, J. D. (2009). Comparison of the Effectiveness of Three Manual Physical Therapy Techniques in a Subgroup of Patients With Low Back Pain Who Satisfy a Clinical Prediction Rule A Randomized Clinical Trial. *Spine (Phila Pa 1976)*, *34*(25), 2720–2729. https://doi.org/10.1097/BRS.0b013e3181b48809

Eklund, A., Bergström, G., Bodin, L. & Axén, I. (2016). Do psychological and behavioral factors classified by the West Haven-Yale Multidimensional Pain Inventory (Swedish version) predict the early clinical course of low back pain in patients receiving chiropractic care? *BMC Musculoskeletal Disorders*, *17*(75). https://doi.org/10.1186/s12891-016-0933-y

Eklund, A., De Carvalho, D., Pagé, I., Wong, A., Johansson, M. S., Pohlman, K. A., Hartvigsen, J. & Swain, M. (2019). Expectations influence treatment outcomes in patients with low back pain. A secondary analysis of data from a randomized clinical trial. *European Journal of Pain (United Kingdom)*, *23*(7), 1378–1389. https://doi.org/10.1002/ejp.1407

Field, J. R. & Newell, D. (2012). Relationship between STarT Back Screening Tool and prognosis for low back pain patients receiving spinal manipulative therapy. *Chiropractic and Manual*

Therapies, *20*(17). https://doi.org/10.1186/2045-709X-20-17

Field, J. R., Newell, D. & McCarthy, P. W. (2010). Preliminary study into the components of the fear-avoidance model of LBP: Change after an initial chiropractic visit and influence on outcome. *Chiropractic and Osteopathy*, *18*, 1–9. https://doi.org/10.1186/1746-1340-18-21

Flynn, T., Fritz, J., Whitman, J., Wainner, R., Magel, J., Rendeiro, D., Butler, B., Garber, M. & Allison, S. (2002). A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine*, *27*(24), 2835–2843. https://doi.org/10.1097/00007632-200212150-00021

Foster, N. E., Anema, J. R., Cherkin, D., Chou, R., Cohen, S. P., Gross, D. P., Ferreira, P. H., Fritz, J. M., Koes, B. W., Peul, W., Turner, J. A., Maher, C. G., Buchbinder, R., Hartvigsen, J., Cherkin, D., Foster, N. E., Maher, C. G., Underwood, M., van Tulder, M., … Woolf, A. (2018). Prevention and treatment of low back pain: evidence, challenges, and promising directions. *The Lancet*, *391*(10137), 2368–2383. https://doi.org/10.1016/S0140-6736(18)30489-6

Fritz, J. M., Brennan, G. P., Clifford, S. N., Hunter, S. J. & Thackeray, A. (2006). An Examination of the Reliability of a Classification Algorithm for Subgrouping Patients With Low Back Pain. *Spine*, *31*(1), 77–82. https://doi.org/10.1097/01.brs.0000193898.14803.8a

Fritz, J. M. & Irrgang, J. J. (2001). A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Physical Therapy*, *81*(2), 776–788. https://doi.org/10.1093/ptj/81.2.776

Fritz, J. M., Koppenhaver, S. L., Kawchuk, G. N., Teyhen, D. S., Hebert, J. J. & Childs, J. D. (2011). Preliminary investigation of the mechanisms underlying the effects of manipulation: exploration of a multivariate model including spinal stiffness, multifidus recruitment, and

clinical findings. *Spine*, *36*(21), 1772–1781. https://doi.org/10.1097/BRS.0b013e318216337d

Fritz, J. M., Sharpe, J. A., Lane, E., Santillo, D., Greene, T. & Kawchuk, G. (2018). Optimizing treatment protocols for spinal manipulative therapy: study protocol for a randomized trial. *Trials*, *19*(1), 306. https://doi.org/10.1186/s13063-018-2692-6

Fritz, J. M. & Wainner, R. S. (2001). Examining diagnostic tests: an evidence-based perspective. *Physical Therapy*, *81*(9), 1546–1564. https://doi.org/10.1093/ptj/81.9.1546

Gatchel, R. J. & Mayer, T. G. (2010). Testing minimal clinically important difference: consensus or conundrum? *Spine Journal*, *10*(4), 321–327. https://doi.org/10.1016/j.spinee.2009.10.015

Glattacker, M., Heyduck, K. & Meffert, C. (2013). Illness beliefs and treatment beliefs as predictors of short-term and medium-term outcome in chronic back pain. *Journal of Rehabilitation Medicine*, *45*(3), 268–276. https://doi.org/10.2340/16501977-1104

Hadizadeh, M., Kawchuk, G. N. & Parent, E. (2019). Reliability of a new loaded rolling wheel system for measuring spinal stiffness in asymptomatic participants. *BMC Musculoskeletal Disorders*, *20*(1), 176. https://doi.org/10.1186/s12891-019-2543-y

Harvey, E., Burton, A. K., Moffett, J. K. & Breen, A. (2003). Spinal manipulation for low-back pain: A treatment package agreed by the UK chiropractic, osteopathy and physiotherapy professional associations. *Manual Therapy*, *8*(1), 46–51. https://doi.org/10.1054/math.2002.0472

Hayden, J. A., Wilson, M. N., Riley, R. D., Iles, R., Pincus, T. & Ogilvie, R. (2019). Individual recovery expectations and prognosis of outcomes in non-specific low back pain: prognostic factor review. *Cochrane Database of Systematic Reviews*, *2019*(11). https://doi.org/10.1002/14651858.CD011284.pub2

Hebert, J. J., Koppenhaver, S. L., Magel, J. S. & Fritz, J. M. (2010). The Relationship of Transversus Abdominis and Lumbar Multifidus Activation and Prognostic Factors for Clinical Success With a Stabilization Exercise Program: A Cross-Sectional Study. *Archives of Physical Medicine and Rehabilitation*, *91*(1), 78–85. https://doi.org/10.1016/j.apmr.2009.08.146

Hill, J. C., Dunn, K. M., Lewis, M., Mullis, R., Main, C. J., Foster, N. E. & Hay, E. M. (2008). A primary care back pain screening tool: Identifying patient subgroups for initial treatment. *Arthritis Care and Research*, *59*(5), 632–641. https://doi.org/10.1002/art.23563

Iben, A., Rosenbaum, A., Röbech, R., Wren, T. & Leboeuf-Yde, C. (2002). Can patient reactions to the first chiropractic treatment predict early favorable treatment outcome in persistent low back pain? *J Manipulative Physiol Ther 2002, 25*(7), 450–454. https://doi.org/10.1067/mmt.2002.126473.

Iles, R. A., Davidson, M., Taylor, N. F. & O'Halloran, P. (2009). Systematic review of the ability of recovery expectations to predict outcomes in non-chronic non-specific low back pain. *Journal of Occupational Rehabilitation*, *19*(1), 25–40. https://doi.org/10.1007/s10926-008-9161-0

Kiesel, K. B., Uhl, T. L., Underwood, F. B., Rodd, D. W. & Nitz, A. J. (2007). Measurement of lumbar multifidus muscle contraction with rehabilitative ultrasound imaging. *Manual Therapy*, *12*(2), 161–166. https://doi.org/10.1016/j.math.2006.06.011

Kongsted, A., Vach, W., Axø, M., Bech, R. N. & Hestbaek, L. (2014). Expectation of recovery from low back pain: A longitudinal cohort study investigating patient characteristics related to expectations and the association between expectations and 3-month outcome. *Spine*, *39*(1), 81–90. https://doi.org/10.1097/BRS.0000000000000059

Koppenhaver, S. L., Fritz, J. M., Hebert, J. J., Kawchuk, G. N., Childs, J. D., Parent, E. C., Gill, N. W. & Teyhen, D. S. (2011). Association between changes in abdominal and lumbar multifidus muscle thickness and clinical improvement after spinal manipulation. *Journal of Orthopaedic and Sports Physical Therapy*, *41*(6), 389–399. https://doi.org/10.2519/jospt.2011.3632

Langworthy, J. M. & Breen, A. C. (2007). Psychosocial factors and their predictive value in chiropractic patients with low back pain: A prospective inception cohort study. *Chiropractic and Osteopathy*, *15*(5), 1–7. https://doi.org/10.1186/1746-1340-15-5

Leboeuf-Yde, C., Axén, I., Jones, J. J., Rosenbaum, A., Løvgren, P. W., Halasz, L. & Larsen, K. (2005). The Nordic Back Pain Subpopulation Program: The long-term outcome pattern in patients with low back pain treated by chiropractors in Sweden. *Journal of Manipulative and Physiological Therapeutics*, *28*(7), 472–478. https://doi.org/10.1016/j.jmpt.2005.07.003

Leboeuf-Yde, C., Grønstvedt, A., Borge, J. A., Lothe, J., Magnesen, E., Nilsson, Ø., Røsok, G., Stig, L. C. & Larsen, K. (2004). The nordic back pain subpopulation program: Demographic and clinical predictors for outcome in patients receiving chiropractic treatment for persistent low-back pain. *Journal of Manipulative and Physiological Therapeutics*, *27*(8), 493–502. https://doi.org/10.1016/j.jmpt.2004.08.001

Leboeuf-yde, C. & Larsen, K. (2005). The Nordic back pain subpopulation program: A I-year prospective multicenter study of outcome of presistent low-back pain in chiropractic patients. *Journal of Manipulative & Physiological Therapeutics*, *28*(2), 90–96. https://doi.org/10.1016/j.jmpt.2005.01.0lO

Leboeuf-Yde, C., Rosenbaum, A., Axén, I., Lövgren, P. W., Jørgensen, K., Halasz, L., Eklund, A. & Wedderkopp, N. (2009). The Nordic Subpopulation Research Programme: Prediction of

treatment outcome in patients with low back pain treated by chiropractors - does the psychological profile matter? *Chiropractic and Osteopathy*, *17*, 1–13. https://doi.org/10.1186/1746-1340-17-14

Magee, D. J. (2007). *Orthopedic Physical Assessment; Chapter 9-Lumbar Spine* (Fifth Edit).

Maher, C. G., Simmonds, M. & Adams, R. (1998). Therapists' conceptualization and characterization of the clinical concept of spinal stiffness. *Physical Therapy*, *78*(3), 289–300. https://doi.org/10.1093/ptj/78.3.289

Maigne, R. (1996). *Diagnosis and Treatment of Pain of Vertebral Origin. A Manual Medicine Approach*. Baltimore: williams & wilkins.

Maitland, G., Hengeveld, E., Banks, K. & English, K. (2005). *Maitland's vertebral manipulation* (7th ed.). Philadelphia: Elsevier.

Malmqvist, S., Leboeuf-Yde, C., Ahola, T., Andersson, O., Ekström, K., Pekkarinen, H., Turpeinen, M. & Wedderkopp, N. (2008). The Nordic back pain subpopulation program: Predicting outcome among chiropractic patients in Finland. *Chiropractic and Osteopathy*, *16*, 1–12. https://doi.org/10.1186/1746-1340-16-13

McCormick, K., Salcedo, J., Peck, J. & Wheeler, A. (2017). *SPSS Statistics for Data Analysis and Visualization.* (1 St). Wiley Publishing.

Metcalfe, C. J. & Moffett, J. A. K. (2005). Do patients' expectations of physiotherapy affect treatment outcome? Part 1- Baseline data. *International Journal of Therapy and Rehabilitation*, *12*(2), 55–62. https://doi.org/10.12968/ijtr.2005.12.2.17456

Meulman, J. (2000). Optimal scaling methods for multivariate categorical data analysis. *Leiden: Leiden University*.

Myers, S. S., Phillips, R. S., Davis, R. B., Cherkin, D. C., Legedza, A., Kaptchuk, T. J., Hrbek, A.,

Buring, J. E., Post, D., Connelly, M. T. & Eisenberg, D. M. (2008). Patient expectations as predictors of outcome in patients with acute low back pain. *Journal of General Internal Medicine*, *23*(2), 148–153. https://doi.org/10.1007/s11606-007-0460-5

Newell, D. & Field, J. (2007). Who will get better? Predicting clinical outcomes in a chiropractic practice. *Clinical Chiropractic*, *10*(4), 179–186. https://doi.org/10.1016/j.clch.2007.10.003

Ostelo, R. W. J. G., Deyo, R. A., Stratford, P., Waddell, G., Croft, P., Von Korff, M., Bouter, L. M. & De Vet, H. C. (2008). Interpreting change scores for pain and functional status in low back pain: Towards international consensus regarding minimal important change. *Spine*, *33*(1), 90–94. https://doi.org/10.1097/BRS.0b013e31815e3a10

Peterson, C. K., Bolton, J. & Humphreys, B. K. (2012). Predictors of improvement in patients with acute and chronic low back pain undergoing chiropractic treatment. *Journal of Manipulative and Physiological Therapeutics*, *35*(7), 525–533. https://doi.org/10.1016/j.jmpt.2012.06.003

Skargren, E. I. & Öberg, B. E. (1998). Predictive factors for 1-year outcome of low-back and neck pain in patients treated in primary care: Comparison between the treatment strategies chiropractic and physiotherapy. *Pain*, *77*(2), 201–207. https://doi.org/10.1016/S0304-3959(98)00101-8

Thakral, M., Shi, L., Shmerling, R. H., Bean, J. F. & Leveille, S. G. (2014). A stiff price to pay: Does joint stiffness predict disability in an older population? *Journal of the American Geriatrics Society*, *62*(10), 1891–1899. https://doi.org/10.1111/jgs.13070

Underwood, M. R., Morton, V. & Farrin, A. (2007). Do baseline characteristics predict response to treatment for low back pain? Secondary analysis of the UK BEAM dataset [ISRCTN32683578]. *Rheumatology*, *46*(8), 1297–1302. https://doi.org/10.1093/rheumatology/kem113

Vavrek, D., Haas, M., Neradilek, M. B. & Polissar, N. (2015). Prediction of pain outcomes in a randomized controlled trial of dose-response of spinal manipulation for the care of chronic low back pain. *BMC Musculoskeletal Disorders*, *16*(1), 1–14. https://doi.org/10.1186/s12891-015-0632-0

Waddell, G., Newton, M., Henderson, I., Somerville, D. & Main, C. J. (1993). A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain*, *52*(2), 157–168. https://doi.org/10.1016/0304-3959(93)90127-B

Waddell, G., Somerville, D., Henderson, I. & Newton, M. (1992). Objective Clinical Evaluation of Physical Impairment in Chronic Low Back Pain. *Spine*, *17*(6), 617–628. https://doi.org/10.1097/00007632-199206000-00001

Wong, A. Y. L. & Kawchuk, G. N. (2017). The clinical value of assessing lumbar posteroanterior segmental stiffness: a narrative review of manual and instrumented methods. *PM R*, *9*(8), 816–830. https://doi.org/10.1016/j.pmrj.2016.12.001

Wong, A. Y. L., Parent, E. C., Dhillon, S. S., Prasad, N., Samartzis, D. & Kawchuk, G. N. (2019). Differential patient responses to spinal manipulative therapy and their relation to spinal degeneration and post-treatment changes in disc diffusion. *European Spine Journal*, *28*(2), 259–269. https://doi.org/10.1007/s00586-018-5851-2

Young, A., Swain, M. S., Kawchuk, G. N., Wong, A. Y. L. & Downie, A. S. (2020). The bench-top accuracy of the VerteTrack spinal stiffness assessment device. *Chiropr Man Ther*, *28*(1), 42. https://doi.org/10.1186/s12998-020-00331-8

Chapter 6. Development of a Multivariable Prediction Model for the Timing of Successful Treatment Response in Patients with Low Back Pain: A Secondary RCT Analysis.

Abstract

Objective: Randomized controlled trials for low back pain typically evaluate responder status at the end of the trial, not when the response is achieved or if the response fluctuates during the trial. Our primary objective was to determine when responder status was achieved over the course of a large low back pain trial and if it was sustained. A secondary objective was to identify which baseline characteristics were predictors of the response patterns of the participants over time.

Methods: We conducted a secondary analysis of a 241-person randomized controlled trial designed to optimize how exercise is combined with spinal manipulative therapy (ClinicalTrials.gov NCT02868034). Responder status was scored as positive if self-reported Oswestry Disability index scores improved by 30% at 1 week, 4 weeks, 12 weeks or never. The resulting 8 response patterns were grouped in two different ways: the Response Onset (when response was first achieved) had 4 collapsed response patterns and Response Persistence (when response was achieved and if it fluctuated) had 5 collapsed response patterns. Between a group's patterns, differences in participant numbers were evaluated (Chi-square Goodness-of-Fit) as were differences in participant baseline characteristics (Chi-square and Kruskal-Wallis). Regardless of

the treatment received, baseline predictors of group patterns were identified through multinomial logistic regression separately for each response definition.

Results: Participant numbers differed significantly ($p<0.05$) between the Response Onset patterns (1-week 28.6%, 4-weeks 29.9%, 12-weeks 16.2%, and non-responders 25.3%) and between the Response Persistence patterns (1-week 22.8%, 4-weeks 26.6%, 12-weeks 16.2%, fluctuating 9.1% and non-responders 25.3%). Baseline variables that univariately differentiated category membership in Response Onset and Response Persistence groupings included pain frequency, depression, neck/upper back pain history, pain intensity, weight, spinal stiffness with STarT Back scores being specific to Response Onset and ultrasonic muscle thickness measurements being specific to Response Persistence. Regression analysis predicted category membership correctly 46.1% for Response Onset and 39.4% correctly for Response Persistence. Maximum terminal stiffness, pain frequency, and neck/upper back pain history appeared in both regression models with lumbar flexion and predicted success with stabilization exercises appearing only in the Response Onset model.

Conclusion: This study is the first to show that treatment response occurs at different times in a large randomized controlled trial among patients with non-specific LBP and can fluctuate over the course of the trial – as in clinical practice. Our findings showed the response patterns are predictable by baseline characteristics. However, the identified predictors need to be investigated further for their potential to modify response by tailoring the intervention to change these modifiable factors. These observations emphasize the limitations of traditional end-of-trial analysis and suggest that future trial analysis be a dynamic process which considers the realities of when the desired response is attained and if it is sustained.

Keywords: Low back pain, spine, spinal manipulative therapy, exercise, response patterns, responders, predictor of response.

6.1 Introduction

Trajectory modeling is a data-driven approach that investigates treatment response variability without assuming all study participants follow a single trajectory (Axén & Leboeuf-Yde, 2013). This approach has shown that low back pain (LBP) is not experienced by people equally with many studies showing that different back pain trajectories exist (Axén & Leboeuf-Yde, 2013; Kongsted et al., 2016) which can be stable or fluctuate over time with various periods of recovery and relapse. Specifically, this approach focuses on the identification of groups of individuals that share common response patterns over time by reducing within-group variability and increasing between-group heterogeneity (Axén & Leboeuf-Yde, 2013; Kongsted et al., 2016; Vasseljen et al., 2013). This method is able to better describe the recurrent and fluctuating pattern of many painful and complex health conditions that do not have a clinical endpoint including LBP (Axén & Leboeuf-Yde, 2013; Dunn et al., 2011). Therefore, despite high clinical heterogeneity, patients with LBP can be grouped into several distinct recovery trajectory classes (Kongsted et al., 2016).

While we now accept that different LBP trajectories exist (Kongsted et al., 2016), research studies often fail to adapt this reality in their analyses. Specifically, randomized clinical trials (RCTs) often establish a single end point for analysis of outcomes such as pain or disability. Given our knowledge of variation in LBP trajectories, this approach may be problematic should some participants achieve the desired response (for a clinically significant period of time), but then lose that response at the analysis end-point (Dutmer et al., 2020). While some analysis techniques consider when the desired response occurs (time-to-event analysis), they do not consider the duration of the response or if the response fluctuates (In & Lee, 2018).

In this secondary analysis of a large LBP RCT, our objectives were to determine 1) when the desired response was achieved during the trial, 2) if the desired response fluctuated during the trial, 3) if participants with distinct response patterns differ in their baseline characteristics, and 4) if baseline characteristics can be used to predict response patterns. We first hypothesized that patients would respond to treatment at different time points through the study and this response may fluctuate. Second, patients with different response patterns would demonstrate statistically significant differences in their baseline characteristics. We finally anticipated that response patterns could be predicted successfully using patients' baseline characteristics.

6.2 Methods

We performed a secondary analysis of data from an RCT (ClinicalTrials.gov, NCT02868034). The RCT was designed to optimize the addition of exercise to SMT in different combinations for patients with LBP by examining the impact of using co-intervention exercise strategies (Fritz et al., 2021). Findings of this study identified SMT sessions followed by multifidus activating exercises as the optimized SMT protocol for the outcome of disability in patients with LBP (Fritz et al., 2021). The full description of the study protocol is reported elsewhere (Fritz et al., 2018). The original study was approved by the Institutional Review Boards at the University of Alberta, Edmonton, AB, Canada (Pro00067152) and the University of Utah, Salt Lake City, UT, USA (IRB_00092127). Here we briefly summarize the method pertinent to the current analysis.

Participants

Individuals with a report of LBP were recruited from the general population over 24 months (February 2017– January 2019). For participants to be eligible, they had to have an Oswestry Disability Index (ODI) of at least 20% and be between 18 and 60 years of age. Exclusion criteria included the presence of any red flags (i.e., tumor, spinal fracture, infection, metabolic diseases, etc.), signs consistent with nerve root compression arising from the clinical examination (e.g., positive straight leg raise test, muscle weakness involving a major muscle group of the lower extremity, diminished lower extremity muscle stretch reflex, etc.), and currently receiving mind-body or exercise treatment for LBP with a healthcare provider (e.g., chiropractor, physical therapist, massage therapist, etc.). Other exclusion criteria included prior surgery to the lumbosacral spine and current pregnancy. Eligible individuals provided informed consent before their enrollment in the study.

6.2.1 Interventions

Participants completed a baseline assessment consisting of i) demographic and clinical history questionnaires, ii) patient-reported outcomes, iii) a physical examination, and iv) instrumented measurements of lumbar spine stiffness and multifidus muscles thickness. All participants then received 2 SMT sessions over the first week. A follow-up visit, 1 week after the initial visit, was scheduled in which items ii– iv from the baseline assessment session were repeated. Participants were then randomly assigned to one of eight groups to receive 6 additional sessions (or no additional treatment) provided over 3 weeks with varied combinations of additional SMT and exercise co-interventions (multifidus muscle activating exercises and/or spine mobilizing

exercises). Two further follow-up visits were scheduled 4 weeks and 12 weeks after the initial visit to monitor recovery by re-collecting items ii– iv from the baseline assessment session (Figure 6-1). Participants were categorized as responders to treatment if their ODI score improved by 30% from baseline independently at each follow-up visit (Gatchel & Mayer, 2010; Ostelo et al., 2008) (Appendix Q).

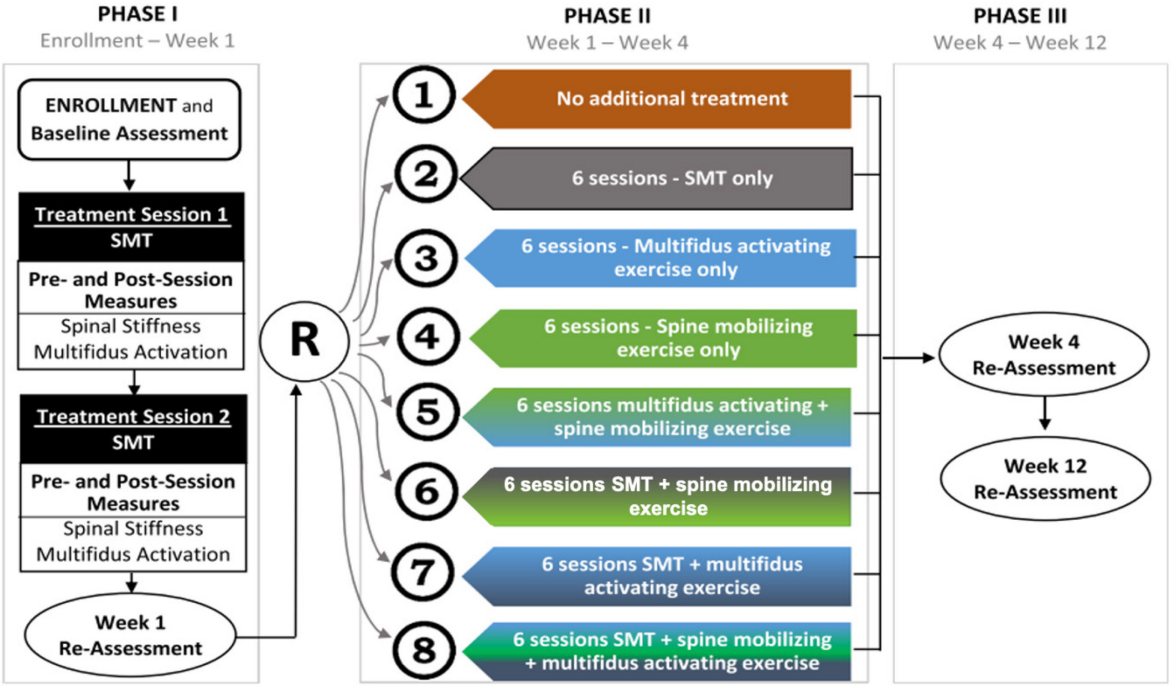


Figure 6-1 Original study design showing three phases and 4 assessment time points.

6.2.1.1 Spinal Manipulative Therapy

A high-velocity, low-amplitude (HVLA) thrust to the anterior superior iliac spine in a posterior and inferior direction (supine position) or to the pelvis in an anterior direction (side-lying position) was performed (Childs et al., 2004; Flynn et al., 2002; Fritz et al., 2011). No difference in outcome

with either the supine or side-posture HVLA technique has been found (Cleland et al., 2009). The clinician could choose the lumbar spine level towards which to direct the manipulation. If cavitation occurred, the SMT treatment was complete. If no cavitation occurred, the participant was repositioned and SMT was performed again. If no cavitation occurred on the second attempt, the clinician manipulated the opposite side. The protocol allowed up to two attempts per side for a successful treatment (Fritz et al., 2018). Substitution with side-posture HVLA was permitted if the supine technique was not possible due to participant preference or comfort.

6.2.1.2 Activating Exercise (ACTex)

The ACTex was a progressive exercise program targeted at the lumbar multifidus muscles. The protocol consisted of isometric multifidus contractions in prone, seated, standing, co-contraction of multifidus and deep abdominal muscles in sitting and standing, and lumbar extensor strengthening exercises (Fritz et al., 2018). Participants performed the isomeric multifidus exercises with an initial dose of 10 repetitions/5 second holds progressing towards 20 repetitions with arm and/or leg lifts was instructed. The general lumbar extensor and multifidus activating exercises began with a dose of 5 repetitions/10 seconds hold and progressed towards 10 repetitions/ 10 seconds hold, 2-3 times per day.

6.2.1.3 Mobilizing Exercise (MOBex)

The MOBex was a repeated active movements progressive program to improve spinal motion and reduce stiffness (Fritz et al., 2018). Participants started with mid-range spinal mobility exercises and progressed into end ranges of either spinal flexion or extension based on the clinician's

direction (McKenzie & May, 2003). Participants were instructed to perform the MOBex with initial dose of 10-20 repetitions and progress towards 40 repetitions throughout the day.

6.2.2 Response Patterns

Participants were classified as responders at each follow-up if they experienced greater than or equal to 30% improvement in ODI relative to baseline. Percentage improvement in ODI score at 1 week was calculated as [baseline ODI score – 1-week ODI score]/ [baseline ODI score *100]. This threshold was based on estimates of minimally meaningful improvement in patients with LBP (Ostelo et al., 2008) and is in line with our previous study (Hadizadeh et al., 2020).

Change in ODI scores were evaluated at three times point in the original study (week 1, week 4, and week 12). Table 6-1 shows the 8 possible response patterns represented as a 3-digit code where the first position represents the 1-week time point, the second position represents the 4-week time point, the third position represents the 12-week time point and "1" codes for a positive response (000, 001, 011, 100, 110, 101, 010, 111.

We then grouped these 8 response patterns in 2 categories: Response Onset (RO) and Response Persistence (RP). The Response Onset grouping with 4 patterns focuses on when a response was first achieved. Specifically, any response pattern of 1XX was considered a 1-week responder, any pattern with a X1X was a 4-week responder while 001 was a 12-week responder, and 000 was non-responder (Table 6-1). The Response Persistence (RP) grouping that has 5 patterns focuses on when response was achieved and if it fluctuated. Therefore, it had 1-week (111), 4-week (011) and

12-week patterns (001) with all other patterns consositng of fluctuating responders who lost their response at some point in the study (010, 100, 110, 101) (Table 6-1).

Table 6-1 Response patterns definitions for different types of analyses

Response timing			**Response pattern**	**Type of analysis**		
Week 1	**Week 4**	**Week 12**	**Pattern**	**Classic End-point at 12 weeks**	**Response Onset**	**Response Persistence**
0	0	x	**001**	Responders	12-week responders	12-week responders
0	x	x	**011**		4-weeks responders	4-weeks responders
x	x	x	**111**		1-week responders	1-week responders
x	0	x	**101**			Fluctuating responders
x	0	0	**100**	Non-responders		
x	x	0	**110**			
0	x	0	**010**		4-weeks responders	
0	0	0	**000**		Non-responders	Non-responders

0= no, 1= yes, reported in the following order: 1week, 4 weeks, 12 weeks

6.2.3 Modelling Variables

We used modelling to determine if the baseline characteristics of trial participants in the RO and RP groupings differed and to predict response groupings from baseline characteristics. To avoid over fitting of a final model, a number of candidate variables were eliminated based on the results from our previous study that did help to identify predicting values of baseline variables for >30% improvement in ODI at 1 week (Hadizadeh et al., 2020). Therefore, all modelling variables were taken from the previous study and consisted of demographic, history, patient-reported outcomes, physical examination, and instrumented variables collected at baseline.

From the demographic and history variables, the following were entered into our analysis: gender, age, weight, height, current LBP episode duration, pain frequency in the past 6 months, comorbid health conditions (history of depression and neck/upper back pain) (Fritz et al., 2018). From the patient-reported outcomes, the following were included: Oswestry disability index (ODI), numeric pain rating scale (NPRS), the 9-item STarT Back Screening Tool, total score of Fear-Avoidance Belief Questionnaire, patients' expectations of LBP outcomes specifically related to medications, rest, manipulation, strengthening, aerobic, and range of motion exercises (a summed value of 6 variables using a 5-point Likert scale). Physical examination findings included assessment of lumbar spine ROM (flexion, extension, worst lateral flexion (minimum between right and left)), and hip ROM (worst hip internal rotation, worst straight leg raise) as measured by inclinometer (Fritz & Piva, 2003), centralization phenomenon (a change in the pain location from a distal or peripheral to a more proximal or central spinal position (Werneke & Hart, 2001) in response to testing repeated lumbar flexion, extension, and lateral flexion), prone instability test (Hicks et al., 2003; McGill et al., 1999), and aberrant movements during ROM (Hicks et al., 2003).

Four variables were used including the prone instability test, aberrant movements during ROM, age, and average SLR that were shown previously to determine predicted success status on the clinical prediction rule for the outcome of lumbar stabilization exercises (Hicks et al., 2005). Instrumented measurements included multifidus muscles thickness as measured by ultrasound imaging as well as global and terminal stiffness measured by VerteTrack at lumbar region (Hadizadeh et al., 2019, 2021). The analysis of ultrasound measures was limited to the minimum and maximum of four variables including multifidus thickness at rest and contracted among the right and left measurements at the L4-L5 and L5-S1 segments. The minimum and maximum

stiffness values from L1 to L5 were selected for the global and terminal spinal stiffness variables. Both spinal stiffness and multifidus muscle thickness variables were normalized by participant's weight to avoid right skewness and allow combination with other participants' data in the modelling.

6.2.4 Statistical Analysis

Sample size calculation was based on previous work in similar patient populations (Fritz et al., 2011). A more detailed explanation of sample size assumptions is provided in the protocol publication (Fritz et al., 2018).

The number of responders was compared within the RO and RP patterns using the Chi-Square Goodness-of-Fit Test. The Chi-square test of independence was performed to compare the baseline values of the categorical variables within the RO and RP groupings. In case data did not meet the assumption of having at least 80% of the cells with an expected count over 5 for the Chi-square test, Fisher's exact test was performed. As the data failed tests for normality of distribution, the Kruskal-Wallis H test was used to compare the non-parametric continuous variables within the RO and RP groupings. Results were considered statistically significant if p values were less than 0.05.

Thereafter, variables that presented statistically significant differences between response pattern categories (using $p \leq 0.2$) were entered into a separate forward stepwise multinomial logistic regression analysis to identify independent predictors of which response patterns a patient might present within RO and RP groupings, respectively. There were no actions taken to blind assessment of predicted outcomes. Relatively low rates of missing data (13.7%) and the

assumption of data missing at random (Sterne et al., 2009), supported the use of multiple imputation for any missing values of the outcome measure to reduce bias (Mackinnon, 2010; Schafer & Graham, 2002). Statistical analyses were conducted using IBM SPSS version 27.0 (Armonk, New York, USA).

6.3 Results

6.3.1 Response patterns

In total, 241 participants were included in all analyses. There was a significant difference between the number of responders in the RO ($\chi^2(2) = 11.1$, $p = .004$), and RP groupings ($\chi^2(3) = 22.8$, $p < .001$). RO consisted of 28.6% (69/241) 1-week responders, 29.9% (72/241) 4-week responders, 16.2% (39/241) 12-week responders, and 25.3% (61/241) non-responders while RP groupings included 22.8% (55/241) 1-week responders, 26.6% (64/241) 4-week responders, 16.2% (39/241) 12-week responders, 9.1% (22/241) fluctuating responders, and 25.3% (61/241) non-responders (Table 6-2).

6.3.2 Baseline characterization of RO and RP groups

Some of the baseline variables showed statistically significant differences within RO and RP groups (Table 6-3 and Table 6-4). The significant baseline variables differentiating category membership within RO and RP included pain frequency, depression, neck or upper back pain history, weight, NPRS, STarT Back Screening Tool score (only RO), minimum multifidus thickness at rest, maximum multifidus thickness at rest (only RP), maximum multifidus thickness contracted (only RP), minimum and maximum global and terminal stiffness ($p \leq .05$). For example, in RP response pattern groupings (Table 6-4), 1-week responders showed lower weight, less pain

frequency, lower pain intensity, higher multifidus muscle thickness at rest and contracted, and higher terminal lumbar stiffness values compared to other groups. A history of depression was more reported in 12-week responders while most fluctuating responders didn't experience depression. A history of neck or upper back pain was mostly reported in non-responders while most fluctuating responders didn't experience neck or upper back pain. Higher global stiffness values were collected from fluctuating responders while the lower global stiffness values were observed in 12-week responders.

6.3.3 Multinominal logistic regression

Model fit was adequate (RO: $\chi^2(15) = 73.50$, $p < .01$; Pseudo-$R^2_{Nagelkerke} = .28$ and RP: $\chi^2(12) = 63.69$, $p < .01$; Pseudo-$R^2_{Nagelkerke} = .24$). Those with higher LBP frequency were more likely to be responders at 1 week (Table 6-5). Those with a history of neck or upper back pain were more likely to be in the 12-week responder group or be non-responders (Table 6-5). If a patient was predicted to respond to stabilization exercises, then they had a lower probability of being a 12-week responder and therefore less likely to be in the non-responder group (Table 6-5).

For one unit increase in the max terminal stiffness value, the probability of someone becoming a 1-week responder was significantly higher than a 4-week or 12-week responder (Table 6-5). In addition, for one unit increase in the lumbar flexion range of motion, the probability of someone becoming a RO/1-week responder was significantly higher relative to becoming a non-responder (Table 6-5).

Overall, regression analysis predicted RO and RP category membership with 46.1% and 39.4% accuracy, respectively (Table 6-6). These results didn't change after omitting missing values.

Table 6-2 Number of patients in each treatment groups based on their response patterns

Treatment Groups	111	011	001	010	100	101	110	000	Total
	1-week responders	4-week responders	12-week responders	Fluctuating responders				Non-responders	
No treatment	7	4	8	2	2	0	0	7	30
SMT	7	5	2	2	2	0	2	9	29
Multifidus activating exercise	7	9	4	1	0	0	0	9	30
Spine mobilizing exercise	8	8	6	1	0	0	0	8	31
Multifidus activating + Spine mobilizing exercise	4	8	6	1	2	1	1	7	30
SMT+ Multifidus activating exercise	8	13	6	0	1	0	0	2	30
SMT+ Spine mobilizing exercise	7	6	2	0	0	1	1	12	29
SMT+ Spine mobilizing+ Multifidus activating exercise	7	11	5	1	0	1	0	7	32
Total	55	64	39	8	7	3	4	61	241

SMT: Spinal Manipulative Therapy
0= no, 1= yes, reported in the following order: 1week, 4 weeks, 12 weeks.

Table 6-3 Baseline characteristics of patients with different RO treatment response patterns.

Characteristic	Overall sample (n=241, 100%)	1-week responders (n=69, 28.63%)	4-week responders (n=72, 29.88%)	12-week responders (n=39, 16.18%)	Non-responders (n=61, 25.31%)	Kruskal-Wallis H/ χ^2/Fisher
Gender n (%)						
Male	97 (40.2)	30 (43.5)	29 (40.3)	14 (35.9)	24 (39.3)	$X^2(3) = .63$
Female	144 (59.8)	39 (56.5)	43 (59.7)	25 (64.1)	37 (60.7)	p-value= .89
Age (yrs)	39.87 (±11.82)	40.32 (±10.77)	38.93 (±11.60)	38.69 (±12.70)	41.23 (±12.74)	H (3) = 1.57 p-value= .67
Weight (kg)	82.9 (±20.7)	78.4 (±18.9)	86.2 (±20.9)	91.4 (±23.6)	78.8 (±18.3)	H (3) = 11.37 p-value= .01*
Height (cm)	171.0 (±10.4)	169.1 (±10.5)	173.4 (±11.8)	170.9 (±8.5)	170.4 (±9.4)	H (3) = 4.12 p-value= .25
Current LBP episode duration n (%)						
< 6 months	75 (31.1)	25 (36.2)	22 (30.6)	9 (23.1)	19 (31.1)	$X^2(3) = 2.03$
> 6 months	166 (68.9)	44 (63.8)	50 (69.4)	30 (76.9)	42 (68.9)	p-value= .57
LBP symptoms n (%)						
Low back only	100 (41.5)	27 (39.1)	29 (40.3)	16 (41.0)	28 (45.9)	$X^2(6) = 7.51$
Buttock(s)	89 (36.9)	33 (47.8)	26 (36.1)	13 (33.3)	17 (27.9)	p-value= .28
Thigh(s)	52 (21.6)	9 (13.0)	17 (23.6)	10 (25.6)	16 (26.2)	
Pain frequency in the past 6 months n (%)						$X^2(3) = 23.20$
At least half the days	198 (82.2)	45 (65.2)	59 (81.9)	36 (92.3)	58 (95.1)	p-value < .01*
Less than half the days	43 (17.8)	24 (34.8)	13 (18.1)	3 (7.7)	3 (4.9)	φ= .31†
History of depression n (%)						$X^2(3) = 15.27$
I do not have this health condition	157 (65.1)	58 (84.1)	42 (58.3)	22 (56.4)	35 (57.4)	p-value < .01*
I had this condition in the past or I currently have it	84 (34.9)	11 (15.9)	30 (41.7)	17 (43.6)	26 (42.6)	φ= .25†

History of neck or upper back pain n (%) I do not have this health condition I had this condition in the past or I currently have it	 134 (55.6) 107 (44.4)	 48 (69.6) 21 (30.4)	 46 (63.9) 26 (36.1)	 17 (43.6) 22 (56.4)	 23 (37.7) 38 (62.3)	$X^2(3) = 17.65$ p-value < .01* φ= .27†
Oswestry disability index (0-100 scale, the higher score indicates greater disability)	34.17 (±11.80)	34.14 (±12.75)	34.47 (±10.60)	34.05 (±12.52)	33.93 (±11.83)	H (3) = .67 p-value= .88
Numeric pain rating scale (0-10 scale, the higher score indicates more pain)	4.61 (±1.64)	4.20 (±1.73)	4.81 (±1.56)	4.41 (±1.52)	4.95 (±1.64)	H (3) = 10.78 p-value= .01*
Fear Avoidance Beliefs Questionnaire (0-96 scale, the higher score indicates more fear avoidance behaviors)	30.03 (±11.54)	27.96 (±10.38)	30.35 (±10.68)	32.33 (±12.08)	30.54 (±13.21)	H (3) = 2.66 p-value= .45
STarT Back Tool (0-9 scale, the higher score indicates higher risk)	4.34 (±1.93)	3.78 (±1.82)	4.44 (±2.01)	4.49 (±2.00)	4.75 (±1.79)	H (3) = 8.42 p-value= .04*
Patients' expectation score (0-30 scale- the higher values indicate the higher expectations)	22.58 (±3.46)	22.78 (±3.13)	22.19 (±4.24)	22.23 (±2.85)	23.02 (±3.15)	H (3) = 2.62 p-value= .46
Lumbar flexion (°)	90.77 (±24.03)	94.96 (±21.43)	88.35 (±25.43)	95.21 (±23.56)	86.07 (±24.70)	H (3) = 4.97 p-value= .17
Lumbar extension (°)	24.15 (±10.16)	22.83 (±10.42)	25.82 (±10.32)	24.92 (±10.12)	23.18 (±9.59)	H (3) = 4.77 p-value= .19
Worst lateral flexion (°)	23.87 (±9.02)	24.58 (±8.59)	24.10 (±9.69)	23.87 (±10.03)	22.80 (±8.07)	H (3) = 1.63 p-value= .65
Centralization phenomenon Yes No	 67 (27.8) 174 (72.2)	 12 (17.4) 57 (82.6)	 22 (30.6) 50 (69.4)	 11 (28.2) 28 (71.8)	 22 (36.1) 39 (63.9)	 $X^2(3) = 6.08$ p-value= .11
Worst hip internal rotation (°)	28.48 (±11.62)	28.30 (±12.24)	29.04 (±11.23)	28.79 (±13.29)	27.82 (±10.41)	H (3) = 1.37 p-value= .71
Worst straight leg raise (°)	70.22 (±16.13)	69.88 (±12.84)	68.85 (±16.73)	70.67(±20.34)	71.93 (±15.96)	H (3) = 1.71 p-value= .64

Success with CPR for stabilization exercises n (%) Negative Positive	 226 (93.8) 15 (6.2)	 62 (89.9) 7 (10.1)	 66 (91.7) 6 (8.3)	 38 (97.4) 1 (2.6)	 60 (98.4) 1 (1.6)	*p*-value=.15
Minimum multifidus muscle thickness rest (mm/kg)	.0172 (±.0053)	.0186 (±.0060)	.0173 (±.0049)	.0156 (±.0046)	.0165 (±.0049)	H (3) = 8.89 *p*-value= .03*
Maximum multifidus muscle thickness rest (mm/kg)	.0207 (±.0060)	.0218 (±.0068)	.0210 (±.0055)	.0189 (±.0061)	.0202 (±.0054)	H (3) = 6.08 *p*-value= .11
Minimum multifidus muscle thickness contracted (mm/kg)	.0189 (±.0054)	.0200 (±.0059)	.0191 (±.0052)	.0173 (±.0050)	.0186 (±.0050)	H (3) = 6.17 *p*-value= .10
Maximum multifidus muscle thickness contracted (mm/kg)	.0228 (±.0061)	.0239 (±.0067)	.0232 (±.0058)	.0210 (±.0063)	.0223 (±.0054)	H (3) = 5.55 *p*-value= .14
Minimum global lumbar stiffness (N/mm.kg)	.0241 (±.0084)	.0259 (±.0083)	.0229 (±.0078)	.0210 (±.0079)	.0256 (±.0090)	H (3) = 11.04 *p*-value= .01*
Maximum global lumbar stiffness (N/mm.kg)	.0295 (±.0103)	.0320 (±.0106)	.0280 (±.0093)	.0252 (±.0088)	.0310 (±.0109)	H (3) = 13.10 *p*-value < .01*
Minimum terminal lumbar stiffness (N/mm.kg)	.0311 (±.0108)	.0335 (±.0108)	.0297 (±.0100)	.0269 (±.0105)	.0329 (±.0111)	H (3) = 11.90 *p*-value= .01*
Maximum terminal lumbar stiffness (N/mm.kg)	.0379 (±.0132)	.0413 (±.0136)	.0362 (±.0124)	.0323 (±.0117)	.0397 (±.0133)	H (3) = 13.74 *p*-value < .01*

Values are mean ±SD unless otherwise indicated.

φ=Effect size (phi coefficient or Cramer's V).

†A significant difference with an effect size _Cohen's definition of ''small''.

*Significantly different at 0.05 level of confidence.

Table 6-4 Baseline characteristics of patients with different RP treatment response patterns.

Characteristic	1-week responders (n=55, 22.82%)	4-week responders (n=65, 26.97%)	12-week responders (n=39, 16.18%)	Fluctuating responders (n=21, 8.71%)	Non-responders (n=61, 25.31%)	Kruskal-Wallis H/ χ^2/Fisher
Gender n (%)						$X^2(4) = .85$
Male	22 (40.0)	27 (41.5)	14 (35.9)	10 (47.6)	24 (39.3)	p-value= .93
Female	33 (60.0)	38 (58.5)	25 (64.1)	11 (52.4)	37 (60.7)	
Age (yrs)	39.91 (±11.04)	39.11 (±11.32)	38.69 (±12.70)	40.38 (±11.60)	41.23 (±12.74)	H (4) = 1.35 p-value= .85
Weight (kg)	76.7 (±17.3)	87.6 (±20.5)	91.4 (±23.6)	81.0 (±23.3)	78.8 (±18.3)	H (4) = 14.46 p-value= .01*
Height (cm)	168.6 (±10.2)	173.9 (±11.7)	170.9 (±8.5)	170.2 (±11.7)	170.4 (±9.4)	H (4) = 6.05 p-value= .20
Current LBP episode duration n (%)						$X^2(4) = 4.11$
< 6 months	21 (38.2)	22 (33.8)	9 (23.1)	4 (19.0)	19 (31.1)	p-value= .40
> 6 months	34 (61.8)	43 (66.2)	30 (76.9)	17 (81.0)	42 (68.9)	
LBP symptoms n (%)						
Low back only	20 (36.4)	27 (41.5)	16 (41.0)	9 (42.9)	28 (45.9)	$X^2(8) = 12.72$
Buttock(s)	30 (54.5)	21 (32.3)	13 (33.3)	8 (38.1)	17 (27.9)	p-value= .12
Thigh(s)	5 (9.1)	17 (26.2)	10 (25.6)	4 (19.0)	16 (26.2)	
Pain frequency in the past 6 months n (%)						$X^2(4) = 29.31$
At least half the days	33 (60.0)	52 (80.0)	36 (92.3)	19 (90.5)	58 (95.1)	p-value < .01*
Less than half the days	22 (40.0)	13 (20.0)	3 (7.7)	2 (9.5)	3 (4.9)	φ= .35†
History of depression n (%)						$X^2(4) = 15.51$
I do not have this health condition	45 (81.8)	37 (56.9)	22 (56.4)	18 (85.7)	35 (57.4)	p-value < .01*
I had this condition in the past or I currently have it	10 (18.2)	28 (43.1)	17 (43.6)	3 (14.3)	26 (42.6)	φ= .25†

History of neck or upper back pain n (%) I do not have this health condition I had this condition in the past or I currently have it	 38 (69.1) 17 (30.9)	 41 (63.1) 24 (36.9)	 17 (43.6) 22 (56.4)	 15 (71.4) 6 (28.6)	 23 (37.7) 38 (62.3)	$X^2(4) = 17.85$ *p*-value < .01* φ= .27†
Oswestry disability index (0-100 scale, the higher score indicates greater disability)	34.73 (±12.72)	34.77 (±10.77)	34.05 (±12.52)	31.81 (±11.64)	33.93 (±11.83)	H (4) = 2.23 *p*-value= .69
Numeric pain rating scale (0-10 scale, the higher score indicates more pain)	4.19 (±1.72)	4.83 (±1.56)	4.41 (±1.52)	4.38 (±1.72)	4.95 (±1.64)	H (4) = 10.28 *p*-value= .04*
Fear Avoidance Beliefs Questionnaire (0-96 scale, the higher score indicates more fear avoidance behaviors)	26.75 (±9.89)	30.82 (±10.85)	32.33 (±12.08)	30.48 (±10.68)	30.54 (±13.21)	H (4) = 4.76 *p*-value= .31
STarT Back Tool (0-9 scale, the higher score indicates higher risk)	3.85 (±1.77)	4.48 (±2.06)	4.49 (±2.00)	3.71 (±1.90)	4.75 (±1.79)	H (4) = 8.48 *p*-value= .08
Patients' expectation score (0-30 scale- the higher values indicate the higher expectations)	22.75 (±3.20)	22.40 (±4.07)	22.23 (±2.85)	22.05 (±4.06)	23.02 (±3.15)	H (4) = 2.41 *p*-value= .66
Lumbar flexion (°)	95.22 (±21.67)	89.86 (±24.34)	95.21 (±23.56)	87.38 (±26.55)	86.07 (±24.70)	H (4) = 4.38 *p*-value= .36
Lumbar extension (°)	23.47 (±11.16)	26.40 (±10.30)	24.92 (±10.12)	20.33 (±7.39)	23.18 (±9.59)	H (4) = 7.82 *p*-value= .10
Worst lateral flexion (°)	24.91 (±8.18)	24.29 (±9.93)	23.87 (±10.03)	22.95 (±9.24)	22.80 (±8.07)	H (4) = 2.49 *p*-value= .65
Centralization phenomenon Yes No	 11 (20.0) 44 (80.0)	 20 (30.8) 45 (69.2)	 11 (28.2) 28 (71.8)	 3 (14.3) 18 (85.7)	 22 (36.1) 39 (63.9)	$X^2(4) = 5.94$ *p*-value= .20
Worst hip internal rotation (°)	27.27 (±11.43)	28.92 (±11.32)	28.79 (±13.29)	31.67 (±13.44)	27.82 (±10.41)	H (4) = 2.69 *p*-value= .61
Worst straight leg raise (°)	68.44 (±13.04)	68.54 (±17.02)	70.67(±20.34)	74.29 (±11.80)	71.93 (±15.96)	H (4) = 3.92 *p*-value= .42

Success with CPR stabilization exercises n (%) Negative Positive	 48 (87.3) 7 (12.7)	 60 (92.3) 5 (7.7)	 38 (97.4) 1 (2.6)	 20 (95.2) 1 (4.8)	 60 (98.4) 1 (1.6)	*p*-value= .12
Minimum multifidus muscle thickness rest (mm/kg)	.0192 (±.0060)	.0173 (±.0051)	.0156 (±.0046)	.0166 (±.0045)	.0165 (±.0049)	H (4) = 12.61 *p*-value= .01*
Maximum multifidus muscle thickness rest (mm/kg)	.0227 (±.0070)	.0208 (±.0056)	.0189 (±.0061)	.0198 (±.0049)	.0202 (±.0054)	H (4) = 9.71 *p*-value= .05*
Minimum multifidus muscle thickness contracted (mm/kg)	.0206 (±.0060)	.0192 (±.0055)	.0173 (±.0050)	.0181 (±.0043)	.0186 (±.0050)	H (4) = 8.64 *p*-value= .07
Maximum multifidus muscle thickness contracted (mm/kg)	.0248 (±.0067)	.0232 (±.0060)	.0210 (±.0063)	.0212 (±.0048)	.0223 (±.0054)	H (4) = 10.95 *p*-value= .03*
Minimum global lumbar stiffness (N/mm.kg)	.0260 (±.0079)	.0224 (±.0074)	.0210 (±.0079)	.0264 (±.0101)	.0256 (±.0090)	H (4) = 13.20 *p*-value= .01*
Maximum global lumbar stiffness (N/mm.kg)	.0320 (±.0103)	.0274 (±.0091)	.0252 (±.0088)	.0325 (±.0114)	.0310 (±.0109)	H (4) = 15.62 *p*-value < .01*
Minimum terminal lumbar stiffness (N/mm.kg)	.0338 (±.0101)	.0290 (±.0095)	.0269 (±.0105)	.0335 (±.0131)	.0329 (±.0111)	H (4) = 14.55 *p*-value= .01*
Maximum terminal lumbar stiffness (N/mm.kg)	.0416 (±.0131)	.0355 (±.0121)	.0323 (±.0117)	.0409 (±.0151)	.0397 (±.0133)	H (4) = 16.08 *p*-value < .01*

φ=Effect size (phi coefficient or Cramer's V).

†A significant difference with an effect size _Cohen's definition of ''small''.

*Significantly different at 0.05 level of confidence.

Table 6-5 Multinomial logistic regression results showing the association between study variables and RO and RP groupings.

	Predictor	ß	Std. Error	Wald	*P*-Value	Odds ratio ($e^{ß}$)	95% Confidence Interval	
							Lower limit	Upper limit
RO								
4-week responders	Maximum terminal lumbar stiffness (N/mm.kg)	-36.55	14.44	6.41	.011	<.0001	<.0001	<0.0001
	LBP frequency	-0.93	.41	5.14	.023	.39	.18	.88
12-week responders	Maximum terminal lumbar stiffness (N/mm.kg)	-71.27	19.20	13.78	<.001	<.0001	<.0001	<.0001
	LBP frequency	-1.81	.67	7.28	.007	.16	.04	.61
	Neck/upper back pain history	1.27	.46	7.54	.006	3.57	1.44	8.84
Non- responders	Lumbar flexion (°)	-.02	.01	7.05	.008	.98	.96	.99
	Success with CPR stabilization exercises	-2.54	1.18	4.67	.031	.08	.01	.79
	LBP frequency	-2.19	.67	10.69	.001	.11	.03	.42
	Neck/upper back pain history	1.76	.43	16.99	<.001	5.82	2.52	13.44
RP								
4-week responders	Maximum terminal lumbar stiffness (N/mm.kg)	-45.04	15.70	8.23	.004	<.0001	<.0001	<.0001
	LBP frequency	-1.08	.43	6.43	.011	.34	.15	.78
12-week responders	Maximum terminal lumbar stiffness (N/mm.kg)	-72.80	19.76	13.58	<.001	<.0001	<.0001	<.0001
	LBP frequency	-2.11	.68	9.63	.002	.12	.03	.46
	Neck/upper back pain history	1.15	.47	5.97	.015	3.16	1.26	7.94

Fluctuating responders	LBP frequency	-1.90	.80	5.67	.017	.15	.03	.72
Non- responders	LBP frequency	-2.58	.67	14.93	<.001	.08	.02	.28
	Neck/upper back pain history	1.32	.42	9.79	.002	3.74	1.64	8.54

RO: Response Onset, RP: Response Persistence

The reference category is 1-week responders in both the RO and the RP analysis.

LBP frequency (0=At least half the days, 1=Less than half the days)

Neck/upper back pain history (0=I do not have this health condition, 1=I had this condition in the past or I currently have it)

Success with CPR stabilization exercises (0=Negative, 1=Positive)

Table 6-6 The observed and the predicted frequencies for RO and RP groupings by multinomial logistic regression for the final models.

	Observed					
Predicted	**1-week responders**	**4-week responders**	**12-week responders**	**Fluctuating responders**	**Non-responders**	**Overall percentage**
Response Onset						
1-week responders	36	22	7	-	9	30.7%
4-week responders	20	31	12	-	13	31.5%
12-week responders	4	4	8	-	3	7.9%
Non-responders	9	15	12	-	36	29.9%
Percent correct	52.2%	43.1%	20.5%	-	59.0%	**46.1%**
Response Persistence						
1-week responders	30	20	6	5	7	28.2%
4-week responders	13	27	12	10	18	33.2%
12-week responders	2	3	7	0	5	7.1%
Fluctuating responders	0	0	0	0	0	0.0%
Non-responders	10	15	14	6	31	31.5%
Overall percentage	54.5%	41.5%	17.9%	0.0%	50.8%	**39.4%**

6.4 Discussion

The present study used prospective data from a large RCT to identify the existence of distinct trial response patterns in patients with non-specific LBP. Further, we showed these patterns could be used to characterize participants at baseline and to predict their treatment response with modest accuracy.

Specifically, our results showed the response of some RCT participants persist over time (90.9%), while other participants lose that response or see their response fluctuate between response states (9.1%). These results will depend largely on what constitutes a minimal response duration as well as a minimal relapse duration – both of which vary in the literature and are not currently standardized (Stanton et al., 2009).

In this study, there were three time points yielding 8 possible responses. As such, these analytical techniques can produce different results depending on how these 8 groups are collapsed or ignored. For example, a classic end-point discounts attributes of some response patterns as part of the non-responder group (010, 100, 110) while including one fluctuating response pattern in the response group (101). Alternatively, the RO and RP groupings do not ignore any response categories (all 8 are accounted for) although each have their strengths and limitations. The RO grouping of response patterns accounts for response onset which classic end-point analysis does not, but the RO approach does not consider response duration or fluctuation - much like a time to effect analysis (i.e. survival analysis) but with many fewer time points which are typically needed for a true survival analysis. Similarly, the RP grouping of response patterns collapses the fluctuating responses together which may or may not be useful depending on what is defined as legitimate response duration and response relapse duration.

Multiple baseline factors predicted membership in both RO and RP categories. Maximum terminal stiffness, pain frequency, and neck/upper back pain history appeared in both regression models with lumbar flexion and clinical-prediction-rule success appearing only in the RO model. The baseline predictive values of pain frequency (Hadizadeh et al., 2020), neck/upper back pain history (Leboeuf-Yde et al., 2004, 2009), and lumbar hypomobility/stiffness (Childs et al., 2004; Flynn et al., 2002), have been reported in previous research in treatment of LBP with the same direction that we found in the present study. Moreover, the present study indicated that fluctuating response to treatment (fluctuating responders) is not common over 12 weeks, only 9.1%, which is small compared with the second least common response pattern (16.18% 12-week responders). Previous

studies regarding the course of LBP reported fluctuating pain in less than 15 % of the cohorts when outcome measures were monitored monthly (27–30) whereas more frequently (25–35 %) with weekly or fortnightly measures (31–34).

Overall, our model to predict treatment response pattern, was not robust (< 50% of patients were correctly predicted) over a 3-month follow-up period among patients with non-specific LBP. Future studies could consider longer follow-up periods, increased number of time points and using cumulative responder duration as an outcome rather than number of responders.

6.4.1 Strengths and Limitations

Strengths of this study include its relatively large sample size, enrollment from two international sites (Utah and Alberta), and the frequency of follow-up time points. We have addressed all the applicable items recommended on reporting of prognostic studies in the TRIPOD checklist (Table 6-7) (Collins et al., 2015). As this was a secondary analysis, we were limited to the information that was collected for the original study. Data regarding the expectation was originally collected on a Likert scale in 12 categories (patient's expectation on medications, surgery, rest, X-ray, MRI, modalities, traction, manipulation, massage, strengthening, aerobic, and range of motion exercises), but in this study, we summed the scores from 5 relevant categories (patient's expectation on rest, manipulation, strengthening, aerobic, and range of motion exercises) to decrease the number of variables. In addition, some baseline predictors were dichotomized to ease interpretation or to meet the assumptions of the regression model. This might have reduced some precision, but probably did not change our conclusions. Finally, our cohort did not all receive the same treatment, and some treatment arms may have had better response overall or different RO

and RP response pattern; we did not examine whether the predictor of response patterns was related to different treatment groups.

Table 6-7 Completeness of reporting of individual TRIPOD items

Item number	**TRIPOD items**	Completeness of reporting
1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	Yes
2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	Yes
3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	Yes
3b	Specify the objectives, including whether the study describes the development or validation of the model or both.	Yes
4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	Yes
4b	Specify the key study dates, including start of accrual, end of accrual, and, if applicable, end of follow-up.	Yes
5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centers.	Yes
5b	Describe eligibility criteria for participants.	Yes
5c	Give details of treatments received, if relevant.	Yes
6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	Yes
6b	Report any actions to blind assessment of the outcome to be predicted.	Yes
7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	Yes
7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	Yes
8	Explain how the study size was arrived at.	Yes
9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	Yes
10a	Describe how predictors were handled in the analyses.	Yes
10b	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	N/A
10c	For validation, describe how the predictions were calculated.	N/A

10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	N/A
10e	Describe any model updating (e.g., recalibration) arising from the validation, if done.	N/A
11	Provide details on how risk groups were created, if done.	N/A
12	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	N/A
13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	Yes
13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	Yes
13c	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors, and outcome).	N/A
14a	Specify the number of participants and outcome events in each analysis.	Yes
14b	If done, report the unadjusted association between each candidate predictor and outcome.	N/A
15a	Present the full prediction model to allow predictions for individuals (i.e. all regression coefficients, and model intercept or baseline survival at a given time point)	Yes
15b	Explain how to the use the prediction model	N/A
16	Report performance measures (with confidence intervals [CIs]) for the prediction model.	Yes
17	If done, report the results from any model updating (i.e. model specification, model performance).	N/A
18	Discuss any limitations of the study (such as non-representative sample, few events per predictor, missing data).	Yes
19a	For validation, discuss the results with reference to performance in the development data and any other validation data.	N/A
19b	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	Yes
20	Discuss the potential clinical use of the model and implications for future research.	Yes
21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	Yes
22	Give the source of funding and the role of the funders for the present study.	Yes

6.5 Conclusion

Trial response patterns were used in the current study for the first time. Our data provided new information on treatment response patterns among patients with non-specific LBP. This shows that response to treatment differs among RCT participants and raises questions of how these changes are accommodated in research versus clinical practice. The current study showed the response patterns were predictable based on baseline characteristics. However, the identified predictors need to be investigated further for their potential to modify treatment response by tailoring interventions to change these modifiable factors. Our findings also challenge the traditional idea of end-of-trial analysis and suggests that it is important to evaluate LBP patients at multiple time points in a study and consider analysis at each time point as the results may be significantly different.

Acknowledgments

The authors would like to thank Dr. Randy Vollrath, Dr. Shannon Wandler, Dr. Moe Gebara, Dr. Lacina Barsalou, Lynne Wong, and Erica Marr for their assistance during the data collection phase. We also acknowledge the National Institutes of Health for funding the original study and for their research funding awarded to MH.

6.6 References

Axén, I. & Leboeuf-Yde, C. (2013). Trajectories of low back pain. *Best Practice and Research: Clinical Rheumatology*, *27*(5), 601–612. https://doi.org/10.1016/j.berh.2013.10.004

Chen, Y., Campbell, P., Strauss, V. Y., Foster, N. E., Jordan, K. P. & Dunn, K. M. (2018). Trajectories and predictors of the long-term course of low back pain: Cohort study with 5-year follow-up. *Pain*, *159*(2), 252–260. https://doi.org/10.1097/j.pain.0000000000001097

Childs, J. D., Fritz, J. M., Flynn, T. W., Irrgang, J. J., Johnson, K. K., Majkowski, G. R. & Delitto, A. (2004). A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: A validation study. *Annals of Internal Medicine*, *141*(12), 920–928. https://doi.org/10.7326/0003-4819-141-12-200412210-00008

Cleland, J. A., Fritz, J. M., Kulig, K., Davenport, T. E., Eberhart, S., Magel, J. & Childs, J. D. (2009). Comparison of the Effectiveness of Three Manual Physical Therapy Techniques in a Subgroup of Patients With Low Back Pain Who Satisfy a Clinical Prediction Rule A Randomized Clinical Trial. *Spine (Phila Pa 1976)*, *34*(25), 2720–2729. https://doi.org/10.1097/BRS.0b013e3181b48809

Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. G. M. (2015). Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. *British Journal of Surgery*, *102*(3), 148–158. https://doi.org/10.1002/bjs.9736

Dunn, K. M., Campbell, P. & Jordan, K. P. (2013). Long-term trajectories of back pain: Cohort study with 7-year follow-up. *BMJ Open*, *3*(12), 1–7. https://doi.org/10.1136/bmjopen-2013-003838

Dunn, K. M., Jordan, K. & Croft, P. R. (2006). Characterizing the course of low back pain: A

latent class analysis. *American Journal of Epidemiology*, *163*(8), 754–761. https://doi.org/10.1093/aje/kwj100

Dunn, K. M., Jordan, K. P., Mancl, L., Drangsholt, M. T. & Le Resche, L. (2011). Trajectories of pain in adolescents: A prospective cohort study. *Pain*, *152*(1), 66–73. https://doi.org/10.1016/j.pain.2010.09.006

Dutmer, A. L., Schiphorst Preuper, H. R., Stewart, R. E., Soer, R., Reneman, M. F. & Wolff, A. P. (2020). Trajectories of Disability and Low Back Pain Impact: 2-year Follow-up of the Groningen Spine Cohort. *Spine*, *45*(23), 1649–1660. https://doi.org/10.1097/BRS.0000000000003647

Flynn, T., Fritz, J., Whitman, J., Wainner, R., Magel, J., Rendeiro, D., Butler, B., Garber, M. & Allison, S. (2002). A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine*, *27*(24), 2835–2843. https://doi.org/10.1097/00007632-200212150-00021

Fritz, J. ., Koppenhaver, S. L., Kawchuk, G. N., Teyhen, D. S., Hebert, J. J. & Childs, J. D. (2011). Preliminary investigation of the mechanisms underlying the effects of manipulation: exploration of a multivariate model including spinal stiffness, multifidus recruitment, and clinical findings. *Spine*, *36*(21), 1772–1781. https://doi.org/10.1097/BRS.0b013e318216337d

Fritz, J. M. & Piva, S. R. (2003). Physical Impairment Index: Reliability, validity, and responsiveness in patients with acute low back pain. *Spine*, *28*(11), 1189–1194. https://doi.org/10.1097/00007632-200306010-00020

Fritz, J. M., Sharpe, J. A., Lane, E., Santillo, D., Greene, T. & Kawchuk, G. (2018). Optimizing treatment protocols for spinal manipulative therapy: study protocol for a randomized trial.

Trials, *19*(1), 306. https://doi.org/10.1186/s13063-018-2692-6

Fritz, J. M., Sharpe, J., Greene, T., Lane, E., Hadizadeh, M., McFadden, M., Santillo, D., Farley, J., Magel, J., Thackeray, A. & Kawchuk, G. (2021). Optimization of Spinal Manipulative Therapy Protocols: A Factorial Randomized Trial Within a Multiphase Optimization Framework. *Journal of Pain*, *22*(6), 655–668. https://doi.org/10.1016/j.jpain.2020.11.008

Gatchel, R. J. & Mayer, T. G. (2010). Testing minimal clinically important difference: consensus or conundrum? *Spine Journal*, *10*(4), 321–327. https://doi.org/10.1016/j.spinee.2009.10.015

Hadizadeh, M., Kawchuk, G. & French, S. (2021). A consensus approach toward the standardization of spinal stiffness measurement using a loaded rolling wheel device: results of a Delphi study. *BMC Musculoskeletal Disorders*, *22*(1), 436. https://doi.org/10.1186/s12891-021-04313-6

Hadizadeh, M., Kawchuk, G. N. & Parent, E. (2019). Reliability of a new loaded rolling wheel system for measuring spinal stiffness in asymptomatic participants. *BMC Musculoskeletal Disorders*, *20*(1), 176. https://doi.org/10.1186/s12891-019-2543-y

Hadizadeh, M., Kawchuk, G. N., Prasad, N. & Fritz, J. M. (2020). Predicting who responds to spinal manipulative therapy using a short-time frame methodology: Results from a 238-participant study. *PLoS ONE*, *15*(11), 1–23. https://doi.org/10.1371/journal.pone.0242831

Hicks, G. E., Fritz, J. M., Delitto, A. & McGill, S. M. (2005). Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization exercise program. *Archives of Physical Medicine and Rehabilitation*, *86*(9), 1753–1762. https://doi.org/10.1016/j.apmr.2005.03.033

Hicks, G. E., Fritz, J. M., Delitto, A. & Mishock, J. (2003). Interrater Reliability of Clinical Examination Measures for Identification of Lumbar Segmental Instability. *Archives of*

Physical Medicine and Rehabilitation, *84*(12), 1858–1864. https://doi.org/10.1016/S0003-9993(03)00365-4

In, J. & Lee, D. K. (2018). Survival analysis: Part I — analysis of time-to-event. *Korean Journal of Anesthesiology*, *71*(13), 182–191. https://doi.org/10.4097/kja.d.18.00067

Kent, P. & Kongsted, A. (2012). Identifying clinical course patterns in SMS data using cluster analysis. *Chiropractic and Manual Therapies*, *20*, 1–12. https://doi.org/10.1186/2045-709X-20-20

Kongsted, A., Kent, P., Axen, I., Downie, A. S. & Dunn, K. M. (2016). What have we learned from ten years of trajectory research in low back pain? *BMC Musculoskeletal Disorders*, *17*(1). https://doi.org/10.1186/s12891-016-1071-2

Kongsted, A., Kent, P., Hestbaek, L. & Vach, W. (2015). Patients with low back pain had distinct clinical course patterns that were typically neither complete recovery nor constant pain. A latent class analysis of longitudinal data. In *Spine Journal* (Vol. 15, Issue 5, pp. 885–894). https://doi.org/10.1016/j.spinee.2015.02.012

Leboeuf-Yde, C., Grønstvedt, A., Borge, J. A., Lothe, J., Magnesen, E., Nilsson, Ø., Røsok, G., Stig, L. C. & Larsen, K. (2004). The nordic back pain subpopulation program: Demographic and clinical predictors for outcome in patients receiving chiropractic treatment for persistent low-back pain. *Journal of Manipulative and Physiological Therapeutics*, *27*(8), 493–502. https://doi.org/10.1016/j.jmpt.2004.08.001

Leboeuf-Yde, C., Rosenbaum, A., Axén, I., Lövgren, P. W., Jørgensen, K., Halasz, L., Eklund, A. & Wedderkopp, N. (2009). The Nordic Subpopulation Research Programme: Prediction of treatment outcome in patients with low back pain treated by chiropractors - does the psychological profile matter? *Chiropractic and Osteopathy*, *17*, 1–13.

https://doi.org/10.1186/1746-1340-17-14

Macedo, L. G., Maher, C. G., Latimer, J., McAuley, J. H., Hodges, P. W. & Rogers, W. T. (2014). Nature and determinants of the course of chronic low back pain over a 12-month period: A cluster analysis. In *Physical Therapy* (Vol. 94, Issue 2, pp. 210–221). https://doi.org/10.2522/ptj.20120416

Mackinnon, A. (2010). The use and reporting of multiple imputation in medical research - a review. *Journal of Internal Medicine*, *268*(6), 586–593. https://doi.org/10.1111/j.1365-2796.2010.02274.x

McGill, S. M., Childs, A. & Liebenson, C. (1999). Endurance times for low back stabilization exercises: Clinical targets for testing and training from a normal database. *Archives of Physical Medicine and Rehabilitation*, *80*(8), 941–944. https://doi.org/10.1016/S0003-9993(99)90087-4

McKenzie, R. & May, S. (2003). *The Lumbar Spine: Mechanical Diagnosis and Therapy* (2nd ed.). Minneapolis, MN, Orthopedic Physical Therapy Products.

Ostelo, R. W. J. G., Deyo, R. A., Stratford, P., Waddell, G., Croft, P., Von Korff, M., Bouter, L. M. & De Vet, H. C. (2008). Interpreting change scores for pain and functional status in low back pain: Towards international consensus regarding minimal important change. *Spine*, *33*(1), 90–94. https://doi.org/10.1097/BRS.0b013e31815e3a10

Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177. https://doi.org/10.1037/1082-989X.7.2.147

Stanton, T. R., Latimer, J., Maher, C. G. & Hancock, M. (2009). Definitions of recurrence of an episode of low back pain: A systematic review. *Spine*, *34*(9), 316–322. https://doi.org/10.1097/BRS.0b013e318198d073

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M. & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *BMJ*, *339*(7713), 157–160. https://doi.org/10.1136/bmj.b2393

Tamcan, O., Mannion, A. F., Eisenring, C., Horisberger, B., Elfering, A. & Müller, U. (2010). The course of chronic and recurrent low back pain in the general population. *Pain*, *150*(3), 451–457. https://doi.org/10.1016/j.pain.2010.05.019

Vasseljen, O., Woodhouse, A., Bjørngaard, J. H. & Leivseth, L. (2013). Natural course of acute neck and low back pain in the general population: The HUNT study. *Pain*, *154*(8), 1237–1244. https://doi.org/10.1016/j.pain.2013.03.032

Werneke, M. & Hart, D. L. (2001). Centralization Phenomenon as a Prognostic Factor for Chronic Low Back Pain and Disability. *Spine*, *26*(7), 758–765. https://doi.org/10.1097/00007632-200104010-00012

Chapter 7. Discussion

The primary purpose of this dissertation was to introduce and standardize a new spinal stiffness measurement device (VT) to provide data for predicting how patients respond to various interventions for LBP. To meet these goals, four studies were undertaken and three papers were published. The first paper, in the third chapter, determined the within- and between-session reliability of lumbar spine stiffness measurements in asymptomatic participants using VT (Hadizadeh et al., 2019). The second paper, in the fourth chapter, developed a standard protocol for evaluating spinal stiffness and to improve the consistency of this assessment in future studies using the VT (Hadizadeh et al., 2021). The third paper, in the fifth chapter, investigated if the prediction of short-term treatment response can be improved by including lumbar spine stiffness measurements in a large LBP RCT (Hadizadeh et al., 2020). Following this work, in the sixth chapter, another secondary analysis of the same RCT was performed to determine if the prediction of long-term treatment response can be improved by including lumbar spine stiffness measurements.

7.1 Summary and Interpretations

A summary of the studies' findings was presented in Figure 7-1. The three published papers in this dissertation, found in chapters 3-5, introduced and standardized a new spinal stiffness measurement device (VT) and used the spinal stiffness data collected by VT to predict how patients respond to various interventions for LBP. With this in mind, we performed a reliability study that evaluated the test-retest reliability of lumbar spine stiffness measurements in asymptomatic individuals using VT. The results of the reliability study indicated that the VT is a reliable assessment device capable of quantifying load-displacement values continuously over an entire

spinal region. A three-round Delphi study was then conducted and as a result, a consensus-based protocol for measuring spinal stiffness using the VT is now available for operators to follow. In total, the pre-defined consensus threshold was reached for 67.2% of statements.

Of the two secondary analyses, the first analysis showed that early response to treatment was predicted by a model containing the following 8 baseline variables: height, gender, neck/upper back pain, pain frequency, the STarT Back Tool, extension status, patients' expectations about medication and strengthening exercises. The results of this study indicated that spinal stiffness as measured by VT was not an important factor in predicting treatment response in the early stage of treatment. The second, secondary analysis identified 8 response patterns in a large LBP RCT and then grouped them in three different ways: classic end-point analysis having 2 categories (response at 12 weeks), Response Onset (RO) having 4 categories (when response was first achieved) and Response Persistence (RP) having 5 categories (when response was achieved and if it fluctuated). Participant numbers differed significantly ($p<0.05$) for classic endpoint (responders 66.8% and non-responders 33.2%), RO categories (1-week 28.6%, 4-weeks 29.9%, 12-weeks 16.2%, and non-responders 25.3%) and for RP categories (1-week 22.8%, 4-weeks 27.0%, 12-weeks 16.2%, fluctuating 9.1% and non-responders 25.3%). Baseline variables that univariately differentiated category membership included pain frequency, depression, neck/upper back pain history, pain intensity, weight, spinal stiffness with STarT Back scores being specific to RO and ultrasonic muscle thickness measurements being specific to RP. Regression analysis predicted category membership correctly 46.1% in RO and 39.4% correctly in RP. Maximum terminal stiffness, pain frequency, and neck/upper back pain history appeared in both regression models with lumbar flexion and predicted success with stabilization exercises appearing only in the RO model. The

findings showed that treatment response in a large RCT occurs at different times and can fluctuate over the course of the trial – as in clinical practice. These observations emphasize the limitations of the classic end-of-trial analysis and suggest that future trial analysis be a dynamic process that considers the realities of when the desired response is attained and if it is sustained. Furthermore, spinal stiffness as measured by VT was a factor in identifying responders when multiple time point measurements were considered.

Study 1 Chapter 3	The within and between-session reliability of lumbar spine stiffness measures using VT at the maximal tolerable load was excellent ranging from 0.95–1.00 and good to excellent ranging from 0.82–0.93, respectively.
Study 2 Chapter 4	Using a standard Delphi methodology, a consensus-based protocol for measuring spinal stiffness was developed. The pre-defined consensus threshold was reached for 67.2% (123/183) of statements after three rounds of surveys.
Study 3 Chapter 5	A predictive model was developed for treatment response after two application of spinal manipulative therapy with a sensitivity of 57.4% and specificity of 91.2%. The final model included height, gender, neck or upper back pain history, pain frequency, the STarT Back Tool, patients' expectations about medication and strengthening exercises, and extension status.
Study 4 Chapter 6	A novel approach was used that considered when responder status was first achieved during the trial (Response Onset (RO)), as well as if responder status was sustained (Response Persistence (RP)). Baseline variables differentiating category membership in RO and RP included pain frequency, depression, neck/upper back pain history, pain intensity, weight, spinal stiffness with STarT Back scores being specific to Response Onset and ultrasonic muscle thickness measurements being specific to Response Persistence. Regression analysis predicted category membership correctly 46.1% for RO and 39.4% correctly for RP. Maximum terminal stiffness, pain frequency, and neck/upper back pain history appeared in both regression models with lumbar flexion and predicted success with stabilization exercises appearing only in the Response Onset model.

Figure 7-1 A summary of the studies' results included in the current dissertation.

7.2 Strengths and Limitations

A full discussion on the strengths and limitations for each study of this dissertation can be found in the individual studies (Chapters 3-6). In general, the strength of my doctoral research is the diversity of methodological approaches used to achieve the primary objective of this dissertation, which was to introduce and standardize a new spinal stiffness measurement device (VT) to provide data for predicting how patients respond to various interventions for LBP. Taking into account the complexity of LBP condition, it can be helpful to use multiple approaches, including both qualitative and quantitative study designs. Other strengths of this dissertation include large sample size, use of measures with substantial evidence of reliability and validity, a longitudinal study design with multivariable analyses and the use of well-known and recognized first and second-line recommended interventions in the management of LBP (Foster et al., 2018) that facilitate the dissemination and clinical application of findings.

There are several considerations in generalizing findings from this research. For instance, the findings have limited generalizability to other populations as the recruited participants for the first study (chapter 3) were healthy individuals (students at the University of Alberta) and for the third and fourth studies (chapters 5 and 6) were non-specific LBP patients. Also, generalizability to other environments (e.g., clinical settings) should be done with caution as participants were tested in two university laboratories. Therefore, the findings may not be representative of all patients with LBP in any environment. Another limitation is our first model did not include the stiffness measurements as a significant predictor of early outcome. One explanation is the original study was designed to assess therapeutic effects in a wide range of participants, so it did not restrict

enrollment to a specific duration of low back pain. Stiffness might have been significant if we had classified our sample in terms of pain duration.

7.3 Implications

Measurement tools (both the subjective and objective measurements) play an important role in research, clinical practice, and health assessment (Souza et al., 2017). Studies on the quality of these tools provide evidence of how the measurement properties were assessed, helping the researcher choose the best tool to use (Souza et al., 2017). Nowadays, a growing number of measurement tools that assess spinal stiffness are available to be used in research. Although many instruments have been created, many of them have not been adequately standardized. Researchers have to carefully choose the reliable and accurate tool, in order to ensure the quality of their results. In the current dissertation, a reliable (Hadizadeh et al., 2019) and accurate (Young, 2019) spinal stiffness measurement device (VT) was introduced. A standard protocol was then developed for the use of the VT device and it's available for all researchers in spine biomechanics and back pain field (Hadizadeh et al., 2021). This standard protocol will facilitate comparability among studies, formal pooling of data, and an increase in large, multicenter studies. Also, it would simplify the process of designing and reviewing research proposals that employ the VT device as well as their manuscripts and publications. Therefore, the knowledge gained from chapters 3 and 4 of this dissertation is intended to provide a foundation for a comprehensive protocol for quantifying stiffness in spine research.

Furthermore, the search for factors that can predict clinical response to the treatment of LBP is a task of substantial practical importance. Identifying patients most likely to respond to a treatment

could improve clinical efficiency and resource utilization (Childs et al., 2004). While the findings suggest that the probability of response to a specific treatment varies significantly between individuals with LBP and can be predicted using the simple prediction models, this does not indicate the proper management. The predicted probabilities of response to treatment at the time points described in chapters 5 and 6 (i.e., week 1, 4 and 12) are important and can be used in shared decision making between patients and clinicians. For example, a patient with the predicted likelihood of favourable response at weeks 1 and 4 may decide to continue minimal care and avoid unnecessary investigations rather than receive additional interventions. While patients who fit the criteria of potential non-responders by 3-months may be more likely to decide to receive more intensive multidisciplinary interventions.

In addition, the current dissertation recommended new analysis methods in identifying response patterns that challenge the traditional idea of end-of-trial analysis and have a potential to make a significant change in response analysis of LBP patients. As such, the findings of prior (and future) studies may differ depending on what time point researchers use to analyze the results of the trial. Baed on the results from the last study (chapter 6), we recommend the Response Persistence analysis as it considers the fluctuating responder group. These findings also suggest that clinicians should view and communicate information about different possible response patterns to provide patients with a realistic view of their condition.

7.4 Future directions

Back pain researchers should strive to use innovative approaches and multiple methodologies, when appropriate to enhance the understanding of this complex condition. This includes the

application of reliable and standard methods in measuring outcome variables such as VT as it was described in Chapters 3 and 4. Mechanical spinal stiffness devices currently have limited utility in clinical practice until further research can determine the appropriateness of these measurements as an outcome measure to guide clinical decision making. This requires a critical review of their psychometric properties, interpretability, and burden of administration. As such, future studies should explore validity and responsiveness of the VT device before any possible clinical recommendation is made. Independent investigations are also required to determine the performance of the VT in other spinal regions and in clinical populations of varying ages and BMIs to determine the true value of this device in these scenarios.

The findings described in Chapter 6 impose important limitations on previous treatment response investigations by only focusing on assessment at the end of the trials. Although using this analysis method gives researchers insights on how effective treatment could be, they miss patients' response patterns through the course of the treatment. This would likely create artificial assumptions about participants' response to a specific treatment. Therefore, understanding patients' response patterns should become more of a priority in back pain research as it has the potential for the development of rehabilitation goals and related outcomes. Consequently, additional response analysis studies are needed to use response pattern and responder duration as recommended in this dissertation to evaluate the efficacy of treatments by when the desired response is attained and if it is sustained. While we presented the first evidence of fluctuating response of LBP RCT participants that challenged the traditional idea of end-of-trial analysis, the detailed implications of this approach in research and clinical practice remain to be answered (e.g., when responder status should be determined during the trial, what would be the best criteria to consider a treatment effective, how

long and/or how often follow-up assessments would be needed to reach the desired response pattern, and whether the response pattern is different in different populations (i.e., patients with acute, subacute, and chronic LBP)). This highlights the need for research to determine whether successful treatment can shift a patient's response pattern to a more optimal one (e.g., a pattern with more and longer positive response). This knowledge would inform a better understanding of what realistic treatment goals in LBP might be. Future RCTs aiming to evaluate treatment response pattern may need to include more frequent assessments of the outcomes (every week, for example) to allow adequate timepoints with equal intervals between the assessments. New technologies including online registries and smartphone apps may provide possibilities to increase the frequency of measurements in larger trials. Additionally, the models developed in chapters 5 and 6 for predicting short-term and long-term responses must be prospectively validated in separate LBP populations across multiple centers before being recommended for clinical implementation (Guyatt et al., 2002).

7.5 Conclusions

Low back pain is a prevalent and challenging condition that has a tremendous impact on the affected individuals, their families, and society. Spinal stiffness assessments are frequently performed as part of LBP patient evaluation in clinical practice. The current doctoral work provides the foundation for a new technique (VT) capable of measuring spinal stiffness continuously over an entire spinal region to study this biomechanical parameter in individuals with LBP. The VT device showed reliable measurements quantifying the load-displacement values for within and between-day assessments in asymptomatic participants. Subsequently, a consensus-based protocol for measuring spinal stiffness using the VT was developed for operators to follow.

It was demonstrated that spinal stiffness as measured by VT was not an important factor in predicting treatment response in the short term but contributed to the final model in predicting responders when more frequent assessment time points were considered. In total, this dissertation suggests spinal stiffness measured by VT has predictive values when long-term and multiple time-points assessments are considered. In addition, this work provided new information on treatment response patterns among patients with non-specific LBP. This shows that treatment response occurs at different times and can fluctuate over the course of the trial – as in clinical practice. These observations emphasized the limitations of traditional end-of-trial analysis and suggest that researchers should adapt to a more dynamic analysis process that considers the realities of when the desired response is attained and if it is sustained.

7.6 References

Childs, J. D., Fritz, J. M., Flynn, T. W., Irrgang, J. J., Johnson, K. K., Majkowski, G. R. & Delitto, A. (2004). A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: A validation study. *Annals of Internal Medicine*, *141*(12), 920–928. https://doi.org/10.7326/0003-4819-141-12-200412210-00008

Foster, N. E., Anema, J. R., Cherkin, D., Chou, R., Cohen, S. P., Gross, D. P., Ferreira, P. H., Fritz, J. M., Koes, B. W., Peul, W., Turner, J. A., Maher, C. G., Buchbinder, R., Hartvigsen, J., Cherkin, D., Foster, N. E., Maher, C. G., Underwood, M., van Tulder, M., … Woolf, A. (2018). Prevention and treatment of low back pain: evidence, challenges, and promising directions. *The Lancet*, *391*(10137), 2368–2383. https://doi.org/10.1016/S0140-6736(18)30489-6

Guyatt, G., Rennie, D., Meade, M. O. & Cook, D. J. (2002). Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. In *Chicago, IL: American Medical Association*.

Hadizadeh, M., Kawchuk, G. & French, S. (2021). A consensus approach toward the standardization of spinal stiffness measurement using a loaded rolling wheel device: results of a Delphi study. *BMC Musculoskeletal Disorders*, *22*(1), 436. https://doi.org/10.1186/s12891-021-04313-6

Hadizadeh, M., Kawchuk, G. N. & Parent, E. (2019). Reliability of a new loaded rolling wheel system for measuring spinal stiffness in asymptomatic participants. *BMC Musculoskeletal Disorders*, *20*(1), 176. https://doi.org/10.1186/s12891-019-2543-y

Hadizadeh, M., Kawchuk, G. N., Prasad, N. & Fritz, J. M. (2020). Predicting who responds to spinal manipulative therapy using a short-time frame methodology: Results from a 238-

participant study. *PLoS ONE*, *15*(11), 1–23. https://doi.org/10.1371/journal.pone.0242831

Souza, A. C. de, Alexandre, N. M. C. & Guirardello, E. de B. (2017). Psychometric properties in instruments evaluation of reliability and validity. *Epidemiologia e Servicos de Saude : Revista Do Sistema Unico de Saude Do Brasil*, *26*(3), 649–659. https://doi.org/10.5123/S1679-49742017000300022

Young, A., Swain, M. S., Kawchuk, G. N., Wong, A. Y. L. & Downie, A. S. (2020). The bench-top accuracy of the VerteTrack spinal stiffness assessment device. *Chiropr Man Ther*, *28*(1), 42. https://doi.org/10.1186/s12998-020-00331-8

Bibliography

Abbott, J. H., Flynn, T. W., Fritz, J. M., Hing, W. A., Reid, D. & Whitman, J. M. (2009). Manual physical assessment of spinal segmental motion: Intent and validity. *Manual Therapy*, *14*(1), 36–44. https://doi.org/10.1016/j.math.2007.09.011

Abbott, J. H., McCane, B., Herbison, P., Moginie, G., Chapple, C. & Hogarty, T. (2005). Lumbar segmental instability: A criterion-related validity study of manual therapy assessment. *BMC Musculoskeletal Disorders*, *6*, 1–10. https://doi.org/10.1186/1471-2474-6-56

Adams, R. (1995). A psychophysical evaluation of manual stiffness discrimination. *Australian Journal of Physiotherapy*, *41*(3), 161–167. https://doi.org/10.1016/S0004-9514(14)60426-8

Akins, R. B., Tolson, H. & Cole, B. R. (2005). Stability of response characteristics of a Delphi panel: application of bootstrap data expansion. *BMC Medical Research Methodology*, *5*, 37. https://doi.org/10.1186/1471-2288-5-37

Angarita-Fonseca, A., Trask, C., Shah, T. & Bath, B. (2019). Stable prevalence of chronic back disorders across gender, age, residence, and physical activity in Canadian adults from 2007 to 2014. *BMC Public Health*, *19*(1), 1–11. https://doi.org/10.1186/s12889-019-7395-8

Asher, A. M., Oleisky, E. R., Pennings, J. S., Khan, I., Sivaganesan, A., Devin, C. J., Bydon, M., Asher, A. L. & Archer, K. R. (2020). Measuring clinically relevant improvement after lumbar spine surgery: is it time for something new? *Spine Journal*, *20*. https://doi.org/10.1016/j.spinee.2020.01.010

Axén, I., Jones, J. J., Rosenbaum, A., Lövgren, P. W., Halasz, L., Larsen, K. & Leboeuf-Yde, C. (2005). The nordic back pain subpopulation program: Validation and improvement of a predictive model for treatment outcome in patients with low back pain receiving chiropractic treatment. *Journal of Manipulative and Physiological Therapeutics*, *28*(6), 381–385.

https://doi.org/10.1016/j.jmpt.2005.06.008

Axén, I. & Leboeuf-Yde, C. (2013). Trajectories of low back pain. *Best Practice and Research: Clinical Rheumatology*, *27*(5), 601–612. https://doi.org/10.1016/j.berh.2013.10.004

Axen, I., Rosenbaum, A., Robech, R., Larsen, K. & Leboeuf-Yde, C. (2005). The Nordic back pain subpopulation program: Can patient reactions to the first chiropractic treatment predict early favorable treatment outcome in nonpersistent low back pain? *Journal of Manipulative and Physiological Therapeutics*, *28*(3), 153–158. https://doi.org/10.1016/j.jmpt.2005.02.007

Bath, B., Trask, C., McCrosky, J. & Lawson, J. (2014). A Biopsychosocial Profile of Adult Canadians with and without Chronic Back Disorders: A Population-Based Analysis of the 2009-2010 Canadian Community Health Surveys. *BioMed Research International*, *2014*, 919621. https://doi.org/10.1155/2014/919621

Baumgart, F. & Cordey, J. (2001). Stiffness - An unknown world of mechanical science? *Injury*, *32*, 14–23. https://doi.org/10.1016/S0020-1383(01)00057-2

Bland, J. M. & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, *8*(2), 135–160. https://doi.org/10.1177/096228029900800204

Bone and Joint Canada. Low Back Pain. (2013). http://www.boneandjointcanada.com/?sec_id=485&msid=3

Bremer, V., Becker, D., Kolovos, S., Funk, B., Van Breda, W., Hoogendoorn, M. & Riper, H. (2018). Predicting therapy success and costs for personalized treatment recommendations using baseline characteristics: Data-driven analysis. *Journal of Medical Internet Research*, *20*(8), 1–11. https://doi.org/10.2196/10275

Brodeur, R. R. & Delre, L. (1999). Stiffness of the Thoracolumbar Spine for Subjects with and

without Low Back Pain. *JNMS - Journal of the Neuromusculoskeletal System*, *7*(4), 127–133.

Brown, B. T., Blacke, A., Carroll, V., Graham, P. L., Kawchuk, G., Downie, A. & Swain, M. (2017). The comfort and safety of a novel rolling mechanical indentation device for the measurement of lumbar trunk stiffness in young adults. *Chiropractic and Manual Therapies*, *25*(21). https://doi.org/10.1186/s12998-017-0153-z

Brunarski, D. (1982). Chiropractic biomechanical evaluations: validity in myofascial low back pain. *J Manipulative Physiol Ther*, *5*(4), 155–161.

Burton, A. K., McClune, T. D., Clarke, R. D. & Main, C. J. (2004). Long-term follow-up of patients with low back pain attending for manipulative care: Outcomes and predictors. *Manual Therapy*, *9*(1), 30–35. https://doi.org/10.1016/S1356-689X(03)00052-3

Burton, A. K., Tillotson, K. M., Main, C. J. & Hollis, S. (1995). Psychosocial predictors of outcome in acute and subchronic low back trouble. *Spine*, *20*(6), 722–728. https://doi.org/10.1097/00007632-199503150-00014

Harvey, D. & Byfieid, D. (1991). Preliminary studies with a mechanical model for the evaluation of spinal motion palpation. *Clin Biomech (Bristol, Avon)*, *6*(2), 79–82. https://doi.org/10.1016/0268-0033(91)90003-9. PMID: 23915479.

Caling, B. & Lee, M. (2001). Effect of direction of applied mobilization force on the posteroanterior response in the lumbar spine. *Journal of Manipulative & Physiological Therapeutics*, *24*(2), 71–78. https://doi.org/10.1067/mmt.2001.112568

Canadian Institute for Health Information (CIHI). *Patient-reported outcome measures (PROMs)*. https://www.cihi.ca/en/patient-reported-outcome-measures-proms

Cecchi, F., Negrini, S., Pasquini, G., Paperini, A., Conti, A., Chiti, M., Zaina, F., Macchi, C. & Molino-Lova, R. (2011). Predictors of functional outcome in patients with chronic low back

pain undergoing back school, individual physiotherapy or spinal manipulation. *European Journal of Physical and Rehabilitation Medicine*, *47*(3), 381–390.

Chen, Y., Campbell, P., Strauss, V. Y., Foster, N. E., Jordan, K. P. & Dunn, K. M. (2018). Trajectories and predictors of the long-term course of low back pain: Cohort study with 5-year follow-up. *Pain*, *159*(2), 252–260. https://doi.org/10.1097/j.pain.0000000000001097

Childs, J.D., Piva, S. R. & Fritz, J. M. (2005). Responsiveness of the Numeric Pain Rating Scale in Patients with Low Back Pain. *Spine*, *30*(1), 1331–1334. https://doi.org/10.1097/01.brs.0000164099.92112.29

Childs, J. D., Fritz, J. M., Flynn, T. W., Irrgang, J. J., Johnson, K. K., Majkowski, G. R. & Delitto, A. (2004). A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: A validation study. *Annals of Internal Medicine*, *141*(12), 920–928. https://doi.org/10.7326/0003-4819-141-12-200412210-00008

Cleland, J. A., Fritz, J. M., Kulig, K., Davenport, T. E., Eberhart, S., Magel, J. & Childs, J. D. (2009). Comparison of the Effectiveness of Three Manual Physical Therapy Techniques in a Subgroup of Patients With Low Back Pain Who Satisfy a Clinical Prediction Rule A Randomized Clinical Trial. *Spine (Phila Pa 1976)*, *34*(25), 2720–2729. https://doi.org/10.1097/BRS.0b013e3181b48809

Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. G. M. (2015). Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The TRIPOD Statement. *British Journal of Surgery*, *102*(3), 148–158. https://doi.org/10.1002/bjs.9736

Colloca, C. J. & Keller, T. S. (2001). Stiffness and neuromuscular reflex response of the human spine to posteroanterior manipulative thrusts in patients with low back pain. *J Manipulative*

Physiol Ther, *24*(8), 489–500. https://doi.org/10.1067/mmt.2001.118209

Colloca, C. J. & Keller, T. S. (2004). Active trunk extensor contributions to dynamic posteroanterior lumbar spinal stiffness. *Journal of Manipulative and Physiological Therapeutics*, *27*(4), 229–237. https://doi.org/10.1016/j.jmpt.2003.11.001

Colloca, C. J., Keller, T. S., Peterson, T. K. & Seltzer, D. E. (2003). Comparison of dynamic posteroanterior spinal stiffness to plain film radiographic images of lumbar dise height. *Journal of Manipulative and Physiological Therapeutics*, *26*(4), 233–241. https://doi.org/10.1016/S0161-4754(03)00005-8

Colloca, C. J. (2010). The Place of Chiropractic Care in the Treatment of Low Back Pain. In M. Szpalski, R. Gunzburg, B. Rydevik, J. Le Huec & H. Mayer (Eds.), *Surgery for Low Back Pain.* (pp. 85–94). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-04547-9

Cruz, E. B., Canhão, H., Fernandes, R., Caeiro, C., Branco, J. C., Rodrigues, A. M., Pimentel-Santos, F., Gomes, L. A., Paiva, S., Pinto, I., Moniz, R. & Nunes, C. (2020). Prognostic indicators for poor outcomes in low back pain patients consulted in primary care. *PLOS ONE*, *15*(3), 1–15. https://doi.org/10.1371/journal.pone.0229265

da Silva, T., Macaskill, P., Mills, K., Maher, C., Williams, C., Lin, C. & Hancock, M. J. (2017). Predicting recovery in patients with acute low back pain: A Clinical Prediction Model. *European Journal of Pain (United Kingdom)*, *21*(4), 716–726. https://doi.org/10.1002/ejp.976

Dagenais, S., Caro, J. & Haldeman, S. (2008). A systematic review of low back pain cost of illness studies in the United States and internationally. *Spine Journal*, *8*(1), 8–20. https://doi.org/10.1016/j.spinee.2007.10.005

Dudek, F. J. (1979). The continuing misinterpretation of the standard error of measurement.

Psychological Bulletin, *86*(2), 335–337. https://doi.org/10.1037/0033-2909.86.2.335

Dunn, K. M., Campbell, P. & Jordan, K. P. (2013). Long-term trajectories of back pain: Cohort study with 7-year follow-up. *BMJ Open*, *3*(12), 1–7. https://doi.org/10.1136/bmjopen-2013-003838

Dunn, K. M., Jordan, K. & Croft, P. R. (2006). Characterizing the course of low back pain: A latent class analysis. *American Journal of Epidemiology*, *163*(8), 754–761. https://doi.org/10.1093/aje/kwj100

Dunn, K. M., Jordan, K. P., Mancl, L., Drangsholt, M. T. & Le Resche, L. (2011). Trajectories of pain in adolescents: A prospective cohort study. *Pain*, *152*(1), 66–73. https://doi.org/10.1016/j.pain.2010.09.006

Dutmer, A. L., Schiphorst Preuper, H. R., Stewart, R. E., Soer, R., Reneman, M. F. & Wolff, A. P. (2020). Trajectories of Disability and Low Back Pain Impact: 2-year Follow-up of the Groningen Spine Cohort. *Spine*, *45*(23), 1649–1660. https://doi.org/10.1097/BRS.0000000000003647

Edmondston, S. J., Allison, G. T., Gregg, C. D., Purden, S. M., Svansson, G. R. & Watson, a. E. (1998). Effect of position on the posteroanterior stiffness of the lumbar spine. *Manual Therapy*, *3*(1), 21–26. https://doi.org/10.1054/math.1998.0312

Eklund, A., Bergström, G., Bodin, L. & Axén, I. (2016). Do psychological and behavioral factors classified by the West Haven-Yale Multidimensional Pain Inventory (Swedish version) predict the early clinical course of low back pain in patients receiving chiropractic care? *BMC Musculoskeletal Disorders*, *17*(75). https://doi.org/10.1186/s12891-016-0933-y

Eklund, A., De Carvalho, D., Pagé, I., Wong, A., Johansson, M. S., Pohlman, K. A., Hartvigsen, J. & Swain, M. (2019). Expectations influence treatment outcomes in patients with low back

pain. A secondary analysis of data from a randomized clinical trial. *European Journal of Pain (United Kingdom)*, *23*(7), 1378–1389. https://doi.org/10.1002/ejp.1407

Fadyl, J. & McPherson, K. (2008). Return to work after injury: A review of evidence regarding expectations and injury perceptions, and their influence on outcome. *Journal of Occupational Rehabilitation*, *18*(4), 362–374. https://doi.org/10.1007/s10926-008-9153-0

Ferrari, S., Vanti, C., Pellizzer, M., Dozza, L., Monticone, M. & Pillastrini, P. (2019). Is there a relationship between self-efficacy, disability, pain and sociodemographic characteristics in chronic low back pain? A multicenter retrospective analysis. *Archives of Physiotherapy*, *9*(1), 1–9. https://doi.org/10.1186/s40945-019-0061-8

Ferreira, M. L., Ferreira, P. H., Latimer, J., Herbert, R. D., Maher, C. & Refshauge, K. (2009). Relationship between spinal stiffness and outcome in patients with chronic low back pain. *Clinical Biomechanics*, *14*, 61–67. https://doi.org/10.1016/j.clinbiomech.2016.03.006

Field, J. R. & Newell, D. (2012). Relationship between STarT Back Screening Tool and prognosis for low back pain patients receiving spinal manipulative therapy. *Chiropractic and Manual Therapies*, *20*(17). https://doi.org/10.1186/2045-709X-20-17

Field, J. R., Newell, D. & McCarthy, P. W. (2010). Preliminary study into the components of the fear-avoidance model of LBP: Change after an initial chiropractic visit and influence on outcome. *Chiropractic and Osteopathy*, *18*, 1–9. https://doi.org/10.1186/1746-1340-18-21

Flynn, T., Fritz, J., Whitman, J., Wainner, R., Magel, J., Rendeiro, D., Butler, B., Garber, M. & Allison, S. (2002). A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine*, *27*(24), 2835–2843. https://doi.org/10.1097/00007632-200212150-00021

Foster, N. E., Anema, J. R., Cherkin, D., Chou, R., Cohen, S. P., Gross, D. P., Ferreira, P. H., Fritz,

J. M., Koes, B. W., Peul, W., Turner, J. A., Maher, C. G., Buchbinder, R., Hartvigsen, J., Cherkin, D., Foster, N. E., Maher, C. G., Underwood, M., van Tulder, M., … Woolf, A. (2018). Prevention and treatment of low back pain: evidence, challenges, and promising directions. *The Lancet*, *391*(10137), 2368–2383. https://doi.org/10.1016/S0140-6736(18)30489-6

Fourney, D. R., Andersson, G., Arnold, P. M., Dettori, J., Cahana, A., Fehlings, M. G., Norvell, D., Samartzis, D. & Chapman, J. R. (2011). Chronic Low Back Pain. *Spine*, *36*(21), S1–S9. https://doi.org/10.1097/BRS.0b013e31822f0a0d

Fritz, J. M. & Irrgang, J. J. (2001). A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Physical Therapy*, *81*(2), 776–788. https://doi.org/10.1093/ptj/81.2.776

Fritz, J. M., Brennan, G. P., Clifford, S. N., Hunter, S. J. & Thackeray, A. (2006). An Examination of the Reliability of a Classification Algorithm for Subgrouping Patients With Low Back Pain. *Spine*, *31*(1), 77–82. https://doi.org/10.1097/01.brs.0000193898.14803.8a

Fritz, J.M., Sharpe, J. A., Lane, E., Santillo, D., Greene, T. & Kawchuk, G. (2018). Optimizing treatment protocols for spinal manipulative therapy: study protocol for a randomized trial. *Trials*, *19*(1), 306. https://doi.org/10.1186/s13063-018-2692-6

Fritz, J. M. & Wainner, R. S. (2001). Examining diagnostic tests: an evidence-based perspective. *Physical Therapy*, *81*(9), 1546–1564. https://doi.org/10.1093/ptj/81.9.1546

Fritz, Julie.M, Koppenhaver, S. L., Kawchuk, G. N., Teyhen, D. S., Hebert, J. J. & Childs, J. D. (2011). Preliminary investigation of the mechanisms underlying the effects of manipulation: exploration of a multivariate model including spinal stiffness, multifidus recruitment, and clinical findings. *Spine*, *36*(21), 1772–1781.

https://doi.org/10.1097/BRS.0b013e318216337d

Fritz, Julie M., Piva, S. R. & Childs, J. D. (2005). Accuracy of the clinical examination to predict radiographic instability of the lumbar spine. *European Spine Journal*, *14*(8), 743–750. https://doi.org/10.1007/s00586-004-0803-4

Fritz, Julie M., Sharpe, J., Greene, T., Lane, E., Hadizadeh, M., McFadden, M., Santillo, D., Farley, J., Magel, J., Thackeray, A. & Kawchuk, G. (2021). Optimization of Spinal Manipulative Therapy Protocols: A Factorial Randomized Trial Within a Multiphase Optimization Framework. *Journal of Pain*, *22*(6), 655–668. https://doi.org/10.1016/j.jpain.2020.11.008

Fritz, Julie M., Whitman, J. M. & Childs, J. D. (2005). Lumbar spine segmental mobility assessment: An examination of validity for determining intervention strategies in patients with low back pain. *Archives of Physical Medicine and Rehabilitation*, *86*(9), 1745–1752. https://doi.org/10.1016/j.apmr.2005.03.028

Fritz, J. M. & Piva, S. R. (2003). Physical Impairment Index: Reliability, validity, and responsiveness in patients with acute low back pain. *Spine*, *28*(11), 1189–1194. https://doi.org/10.1097/00007632-200306010-00020

Gatchel, R. J. & Mayer, T. G. (2010). Testing minimal clinically important difference: consensus or conundrum? *Spine Journal*, *10*(4), 321–327. https://doi.org/10.1016/j.spinee.2009.10.015

George, S. Z., Lentz, T. A., Beneciuk, J. M., Bhavsar, N. A., Mundt, J. M. & Boissoneault, J. (2020). Framework for improving outcome prediction for acute to chronic low back pain transitions. *PAIN Reports*, *5*(2), e809. https://doi.org/10.1097/pr9.0000000000000809

Geurts, J. W., Willems, P. C., Kallewaard, J. W., Van Kleef, M. & Dirksen, C. (2018). The Impact of Chronic Discogenic Low Back Pain: Costs and Patients' Burden. *Pain Research and Management*, *2018*. https://doi.org/10.1155/2018/4696180

Glattacker, M., Heyduck, K. & Meffert, C. (2013). Illness beliefs and treatment beliefs as predictors of short-term and medium-term outcome in chronic back pain. *Journal of Rehabilitation Medicine*, *45*(3), 268–276. https://doi.org/10.2340/16501977-1104

Grieve, G. (1982). Lumbar instability. *Physiotherapy*, *68*, 2–9.

Grieve, G. (1989). *Common vertebral joint problems*. Edinburgh: Churchill Livingstone.

Gross, D. P., Ferrari, R., Russell, A. S., Battié, M. C., Schopflocher, D., Hu, R. W., Waddell, G. & Buchbinder, R. (2006). A population-based survey of back pain beliefs in Canada. *Spine*, *31*(18), 2142–2145. https://doi.org/10.1097/01.brs.0000231771.14965.e4

Guyatt, G., Rennie, D., Meade, M. O. & Cook, D. J. (2002). Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. In *Chicago, IL: American Medical Association*.

Hadizadeh, M., Kawchuk, G. & French, S. (2021). A consensus approach toward the standardization of spinal stiffness measurement using a loaded rolling wheel device: results of a Delphi study. *BMC Musculoskeletal Disorders*, *22*(1), 436. https://doi.org/10.1186/s12891-021-04313-6

Hadizadeh, M., Kawchuk, G. N. & Parent, E. (2019). Reliability of a new loaded rolling wheel system for measuring spinal stiffness in asymptomatic participants. *BMC Musculoskeletal Disorders*, *20*(1), 176. https://doi.org/10.1186/s12891-019-2543-y

Hadizadeh, M., Kawchuk, G. N., Prasad, N. & Fritz, J. M. (2020). Predicting who responds to spinal manipulative therapy using a short-time frame methodology: Results from a 238-participant study. *PLoS ONE*, *15*(11), 1–23. https://doi.org/10.1371/journal.pone.0242831

Harrell, F. E. (2015). Regression modeling strategies with applications to linear models, logistic regression and survival analysis. In *Springer* (2nd ed.).

https://doi.org/10.1198/tech.2003.s158

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N. & Conde, J. G. (2009). Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010

Harsted, S., Nyirö, L., Downie, A., Kawchuk, G. N., O'Neill, S., Holm, L. & Nim, C. G. (2021). Posterior to anterior spinal stiffness measured in a sample of 127 secondary care low back pain patients. *Clinical Biomechanics*, 105408. https://doi.org/10.1016/j.clinbiomech.2021.105408

Hartvigsen, J., Hancock, M. J., Kongsted, A., Louw, Q., Ferreira, M. L., Genevay, S., Hoy, D. & Karppinen, J. (2018). *Series Low back pain 1 What low back pain is and why we need to pay attention*. *391*. https://doi.org/10.1016/S0140-6736(18)30480-X

Harvey, E., Burton, A. K., Moffett, J. K. & Breen, A. (2003). Spinal manipulation for low-back pain: A treatment package agreed by the UK chiropractic, osteopathy and physiotherapy professional associations. *Manual Therapy*, *8*(1), 46–51. https://doi.org/10.1054/math.2002.0472

Hasson, F., Keeney, S. & Mckenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, *32*(4), 1008–1015.

Hayden, J. A., Wilson, M. N., Riley, R. D., Iles, R., Pincus, T. & Ogilvie, R. (2019). Individual recovery expectations and prognosis of outcomes in non-specific low back pain: prognostic factor review. *Cochrane Database of Systematic Reviews*, *2019*(11). https://doi.org/10.1002/14651858.CD011284.pub2

Hebert, J. J., Koppenhaver, S. L., Magel, J. S. & Fritz, J. M. (2010). The Relationship of

Transversus Abdominis and Lumbar Multifidus Activation and Prognostic Factors for Clinical Success With a Stabilization Exercise Program: A Cross-Sectional Study. *Archives of Physical Medicine and Rehabilitation*, *91*(1), 78–85. https://doi.org/10.1016/j.apmr.2009.08.146

Hestbœk, L. & Leboeuf-Yde, C. (2000). Are chiropractic tests for the lumbo-pelvic spine reliable and valid? A systematic critical literature review. *Journal of Manipulative and Physiological Therapeutics*, *23*(4), 258–275. https://doi.org/10.1067/mmt.2000.106097

Hicks, G. E., Fritz, J. M., Delitto, A. & McGill, S. M. (2005). Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization exercise program. *Archives of Physical Medicine and Rehabilitation*, *86*(9), 1753–1762. https://doi.org/10.1016/j.apmr.2005.03.033

Hicks, G. E., Fritz, J. M., Delitto, A. & Mishock, J. (2003). Interrater Reliability of Clinical Examination Measures for Identification of Lumbar Segmental Instability. *Archives of Physical Medicine and Rehabilitation*, *84*(12), 1858–1864. https://doi.org/10.1016/S0003-9993(03)00365-4

Hill, J. C., Whitehurst, D. G. T., Lewis, M., Bryan, S., Dunn, K. M., Foster, N. E., Konstantinou, K., Main, C. J., Mason, E., Somerville, S., Sowden, G., Vohora, K. & Hay, E. M. (2011). Comparison of stratified primary care management for low back pain with current best practice (STarT Back): A randomised controlled trial. *The Lancet*, *378*(9802), 1560–1571. https://doi.org/10.1016/S0140-6736(11)60937-9

Hill, J. C., Dunn, K. M., Lewis, M., Mullis, R., Main, C. J., Foster, N. E. & Hay, E. M. (2008). A primary care back pain screening tool: Identifying patient subgroups for initial treatment. *Arthritis Care and Research*, *59*(5), 632–641. https://doi.org/10.1002/art.23563

Hoy, D., Bain, C., Williams, G., March, L., Brooks, P., Blyth, F., Woolf, A., Vos, T. & Buchbinder, R. (2012). A systematic review of the global prevalence of low back pain. *Arthritis and Rheumatism*, *64*(6), 2028–2037. https://doi.org/10.1002/art.34347

Huijbregts, P. A. (2002). Spinal motion palpation: A review of reliability studies. *Journal of Manual and Manipulative Therapy*, *10*(1), 24–39. https://doi.org/10.1179/106698102792209585

Hurwitz, E. L., Randhawa, K., Yu, H., Côté, P. & Haldeman, S. (2018). The Global Spine Care Initiative: a summary of the global burden of low back and neck pain studies. *European Spine Journal*, *27*(s6), 796–801. https://doi.org/10.1007/s00586-017-5432-9

Iben, A., Rosenbaum, A., Röbech, R., Wren, T. & Leboeuf-Yde, C. (2002). Can patient reactions to the first chiropractic treatment predict early favorable treatment outcome in persistent low back pain? *J Manipulative Physiol Ther*, *25*(7), 450–454. https://doi.org/10.1067/mmt.2002.126473.

Iles, R. A., Davidson, M., Taylor, N. F. & O'Halloran, P. (2009). Systematic review of the ability of recovery expectations to predict outcomes in non-chronic non-specific low back pain. *Journal of Occupational Rehabilitation*, *19*(1), 25–40. https://doi.org/10.1007/s10926-008-9161-0

In, J. & Lee, D. K. (2018). Survival analysis: Part I — analysis of time-to-event. *Korean Journal of Anesthesiology*, *71*(13), 182–191. https://doi.org/10.4097/kja.d.18.00067

Iqbal, S. & Pipon-Young, L. (2009). The Delphi method. *Nursing Research*, *46*(2), 116–118.

Jenson, K., Gemmell, H. & Thiel, H. W. (1993). Motion palpation accuracy using a mechanical spinal model. *European Journal of Chiropractic*, *41*, 67–73.

Jun, P., Pagé, I., Vette, A. & Kawchuk, G. (2020). Potential mechanisms for lumbar spinal stiffness

change following spinal manipulative therapy : a scoping review. *Chiropr Man Therap*, *28*(1), 15. https://doi.org/10.1186/s12998-020-00304-x

Kassebaum, N. J., Arora, M., Barber, R. M., Brown, J., Carter, A., Casey, D. C., Charlson, F. J., Coates, M. M., Coggeshall, M., Cornaby, L., Dandona, L., Dicker, D. J., Erskine, H. E., Ferrari, A. J., Fitzmaurice, C., Foreman, K., Forouzanfar, M. H., Fullman, N., Goldberg, E. M., … Zuhlke, L. J. (2016). GBD 2015 DALYs and HALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy life expectancy (HALE), 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet*, *388*(10053), 1603–1658. https://doi.org/10.1016/S0140-6736(16)31460-X

Katz, J. N. (2006). Lumbar disc disorders and low-back pain: Socioeconomic factors and consequences. *Journal of Bone and Joint Surgery - Series A*, *88*(SUPPL. 2), 21–24. https://doi.org/10.2106/JBJS.E.01273

Kawchuk, G. N., Kaigle, A. M., Holm, S. H., Fauvel, O. R., Ekström, L. & Hansson, T. (2001). The diagnostic performance of vertebral displacement measurements derived from ultrasonic indentation in an in vivo model of degenerative disc disease. *Spine*, *26*(12), 1348–1355. https://doi.org/10.1097/00007632-200106150-00018

Kay, C., Wozniak, E. & Bernstein, J. (2017). Utilization of Health Care Services and Ambulatory Resources Associated with Chronic Noncancer Pain. *Pain Medicine (Malden, Mass.)*, *18*(7), 1236–1246. https://doi.org/10.1093/pm/pnw336

Keller, T. S., Colloca, C. J. & Fuhr, A. W. (1999). Validation of the force and frequency characteristics of the Activator Adjusting Instrument: Effectiveness as a mechanical impedance measurement tool. *Journal of Manipulative and Physiological Therapeutics*,

22(2), 75–86. https://doi.org/10.1016/S0161-4754(99)70111-9

Kenna, C. & Murtagh, J. (1997). *Back Pain and Spinal Manipulation: A Practical Guid* (2nd ed.). Butterworth-Heinemann.

Keeney, S., Hasson, F. & Mckenna, H. (2006). Consulting the oracle : ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing*, *53*(2), 205–212. https://doi.org/10.1111/j.1365-2648.2006.03716.x

Keeney, S., Mckenna, H. & Hasson, F. (2011). *The Delphi technique in nursing and health research*. West Sussex, United Kingdom: Wiley-Blackwell.

Kent, P. & Kongsted, A. (2012). Identifying clinical course patterns in SMS data using cluster analysis. *Chiropractic and Manual Therapies*, *20*, 1–12. https://doi.org/10.1186/2045-709X-20-20

Kiesel, K. B., Uhl, T. L., Underwood, F. B., Rodd, D. W. & Nitz, A. J. (2007). Measurement of lumbar multifidus muscle contraction with rehabilitative ultrasound imaging. *Manual Therapy*, *12*(2), 161–166. https://doi.org/10.1016/j.math.2006.06.011

Kim, L. H., Vail, D., Azad, T. D., Bentley, J. P., Zhang, Y., Ho, A. L., Fatemi, P., Feng, A., Varshneya, K., Desai, M., Veeravagu, A. & Ratliff, J. K. (2019). Expenditures and Health Care Utilization Among Adults With Newly Diagnosed Low Back and Lower Extremity Pain. *JAMA Network Open*, *2*(5), e193676. https://doi.org/10.1001/jamanetworkopen.2019.3676

Koes, B. W., Van Tulder, M. W. & Thomas, S. (2006). Diagnosis and Treatment of Low Back Pain. *Bmj*, *332*(7555), 1430–1434. https://doi.org/10.1136/bmj.332.7555.1430

Kongsted, A., Andersen, C. H., Hansen, M. M. & Hestbaek, L. (2016). Prediction of outcome in patients with low back pain - A prospective cohort study comparing clinicians' predictions with those of the Start Back Tool. *Manual Therapy*, *21*, 120–127.

https://doi.org/10.1016/j.math.2015.06.008

Kongsted, A., Kent, P., Axen, I., Downie, A. S. & Dunn, K. M. (2016). What have we learned from ten years of trajectory research in low back pain? *BMC Musculoskeletal Disorders*, *17*(1). https://doi.org/10.1186/s12891-016-1071-2

Kongsted, A., Vach, W., Axø, M., Bech, R. N. & Hestbaek, L. (2014). Expectation of recovery from low back pain: A longitudinal cohort study investigating patient characteristics related to expectations and the association between expectations and 3-month outcome. *Spine*, *39*(1), 81–90. https://doi.org/10.1097/BRS.0000000000000059

Kongsted, A., Kent, P., Hestbaek, L. & Vach, W. (2015). Patients with low back pain had distinct clinical course patterns that were typically neither complete recovery nor constant pain. A latent class analysis of longitudinal data. *Spine Journal*, *15*(5), 885–894. https://doi.org/10.1016/j.spinee.2015.02.012

Koo, T. K. & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Koppenhaver, S. L., Fritz, J. M., Hebert, J. J., Kawchuk, G. N., Childs, J. D., Parent, E. C., Gill, N. W. & Teyhen, D. S. (2011). Association between changes in abdominal and lumbar multifidus muscle thickness and clinical improvement after spinal manipulation. *Journal of Orthopaedic and Sports Physical Therapy*, *41*(6), 389–399. https://doi.org/10.2519/jospt.2011.3632

Koppenhaver, S. L., Hebert, J. J., Kawchuk, G. N., Childs, J. D., Teyhen, D. S., Croy, T. & Fritz, J. M. (2014). Criterion validity of manual assessment of spinal stiffness. *Manual Therapy*, *19*(6), 589–594. https://doi.org/10.1016/j.math.2014.06.001

Landel, R., Kulig, K., Fredericson, M., Li, B. & Powers, C. M. (2008). Intertester reliability and validity of motion assessments during lumbar spine accessory motion testing. *Phys Ther*, *88*(1), 43–49. https://doi.org/10.2522/ptj.20060179

Langworthy, J. M. & Breen, A. C. (2007). Psychosocial factors and their predictive value in chiropractic patients with low back pain: A prospective inception cohort study. *Chiropractic and Osteopathy*, *15*(5), 1–7. https://doi.org/10.1186/1746-1340-15-5

Latimer, J., Goodsel, M. M., Lee, M., Maher, C. G., Wilkinson, B. N. & Moran, C. C. (1996). Evaluation of a new device for measuring responses to posteroanterior forces in a patient population, part 1: Reliability testing. *Physical Therapy*, *76*(2), 158–165. https://doi.org/10.1093/ptj/76.2.158

Latimer, J., Lee, M. & Adams, R. D. (1998). The effects of high and low loading forces on measured values of lumbar stiffness. *Journal of Manipulative and Physiological Therapeutics*, *21*(3), 157–163.

Latimer, J., Lee, M., Adams, R. & Moran, C. (1996). An Investigation of the Relationship between Low Back Pain and Lumbar Posteroanterior Stiffness. *Journal of Manipulative & Physiological Therapeutics*, *19*(9), 587–591.

Leboeuf-Yde, C., Axén, I., Jones, J. J., Rosenbaum, A., Løvgren, P. W., Halasz, L. & Larsen, K. (2005). The Nordic Back Pain Subpopulation Program: The long-term outcome pattern in patients with low back pain treated by chiropractors in Sweden. *Journal of Manipulative and Physiological Therapeutics*, *28*(7), 472–478. https://doi.org/10.1016/j.jmpt.2005.07.003

Leboeuf-Yde, C., Grønstvedt, A., Borge, J. A., Lothe, J., Magnesen, E., Nilsson, Ø., Røsok, G., Stig, L. C. & Larsen, K. (2004). The nordic back pain subpopulation program: Demographic and clinical predictors for outcome in patients receiving chiropractic treatment for persistent

low-back pain. *Journal of Manipulative and Physiological Therapeutics*, *27*(8), 493–502. https://doi.org/10.1016/j.jmpt.2004.08.001

Leboeuf-yde, C. & Larsen, K. (2005). The Nordic back pain subpopulation program: A I-year prospective multicenter study of outcome of presistent low-back pain in chiropractic patients. *Journal of Manipulative & Physiological Therapeutics*, *28*(2), 90–96. https://doi.org/10.1016/j.jmpt.2005.01.0lO

Leboeuf-Yde, C., Rosenbaum, A., Axén, I., Lövgren, P. W., Jørgensen, K., Halasz, L., Eklund, A. & Wedderkopp, N. (2009). The Nordic Subpopulation Research Programme: Prediction of treatment outcome in patients with low back pain treated by chiropractors - does the psychological profile matter? *Chiropractic and Osteopathy*, *17*, 1–13. https://doi.org/10.1186/1746-1340-17-14

Lederer, A.-K., Maly, C., Weinert, T. & Huber, R. (2019). Tissue Stiffness is Not Related to Pain Experience: An Individually Controlled Study in Patients with Chronic Neck and Back Pain. *Evid Based Complement Alternat Med*, *2019*. https://doi.org/10.1155/2019/1907168

Lee, M. & Svensson, N. L. (1990). Measurement of stiffness during simulated spinal physiotherapy. *Clin Phys Physiol Meas*, *11*(3), 201–207. https://doi.org/10.1088/0143-0815/11/3/002

Lee, M. & Svensson, N. L. (1993). Effect of loading frequency on response of the spine to lumbar posteroanterior forces. *J Manipulative Physiol Ther*, *16*, 439–446.

Lim, K. L., Jacobs, P. & Klarenbach, S. (2006). A population-based analysis of healthcare utilization of persons with back disorders: Results from the Canadian Community Health Survey 2000-2001. *Spine*, *31*(2), 212–218. https://doi.org/10.1097/01.brs.0000194773.10461.9f

Macedo, L. G., Maher, C. G., Latimer, J., McAuley, J. H., Hodges, P. W. & Rogers, W. T. (2014). Nature and determinants of the course of chronic low back pain over a 12-month period: A cluster analysis. *Physical Therapy*, *94*(2), 210–221. https://doi.org/10.2522/ptj.20120416

Mackinnon, A. (2010). The use and reporting of multiple imputation in medical research - a review. *Journal of Internal Medicine*, *268*(6), 586–593. https://doi.org/10.1111/j.1365-2796.2010.02274.x

McGill, S. M., Childs, A. & Liebenson, C. (1999). Endurance times for low back stabilization exercises: Clinical targets for testing and training from a normal database. *Archives of Physical Medicine and Rehabilitation*, *80*(8), 941–944. https://doi.org/10.1016/S0003-9993(99)90087-4

McKenzie, R. & May, S. (2003). *The Lumbar Spine: Mechanical Diagnosis and Therapy* (2nd ed.). Minneapolis, MN, Orthopedic Physical Therapy Products.

Magee, D. J. (2007). *Orthopedic Physical Assessment; Chapter 9-Lumbar Spine* (Fifth Edit).

Maher, C. & Adams, R. (1994). Reliability of Pain and Stiffness Assessments in Clinical Manual Lumbar Spine Examination. *Phys Ther*, *74*(9), 801–809. https://doi.org/10.1093/ptj/74.9.801

Maher, C. G., Simmonds, M. & Adams, R. (1998). Therapists' conceptualization and characterization of the clinical concept of spinal stiffness. *Physical Therapy*, *78*(3), 289–300. https://doi.org/10.1093/ptj/78.3.289

Maher, C., Underwood, M. & Buchbinder, R. (2017). Non-specific low back pain. *The Lancet*, *389*(10070), 736–747. https://doi.org/10.1016/S0140-6736(16)30970-9

Maigne, R. (1996). *Diagnosis and Treatment of Pain of Vertebral Origin. A Manual Medicine Approach*. Baltimore: williams & wilkins.

Maitland, G., Hengeveld, E., Banks, K. & English, K. (2005). *Maitland's vertebral manipulation*

(7th ed.). Philadelphia: Elsevier.

Malmqvist, S., Leboeuf-Yde, C., Ahola, T., Andersson, O., Ekström, K., Pekkarinen, H., Turpeinen, M. & Wedderkopp, N. (2008). The Nordic back pain subpopulation program: Predicting outcome among chiropractic patients in Finland. *Chiropractic and Osteopathy*, *16*, 1–12. https://doi.org/10.1186/1746-1340-16-13

McCormick, K., Salcedo, J., Peck, J. & Wheeler, A. (2017). *SPSS Statistics for Data Analysis and Visualization.* (1 St). Wiley Publishing.

McKenzie, R. & May, S. (2003). *The Lumbar Spine: Mechanical Diagnosis and Therapy* (2nd ed.). Minneapolis, MN, Orthopedic Physical Therapy Products.

Mendonça, L., Monteiro-Soares, M. & Azevedo, L. F. (2018). Prediction of clinical outcomes in individuals with chronic low back pain: A protocol for a systematic review with meta-analysis 11 Medical and Health Sciences 1117 Public Health and Health Services. *Systematic Reviews*, *7*(1), 1–6. https://doi.org/10.1186/s13643-018-0818-2

Metcalfe, C. J. & Moffett, J. A. K. (2005). Do patients' expectations of physiotherapy affect treatment outcome? Part 1- Baseline data. *International Journal of Therapy and Rehabilitation*, *12*(2), 55–62. https://doi.org/10.12968/ijtr.2005.12.2.17456

Meulman, J. (2000). Optimal scaling methods for multivariate categorical data analysis. *Leiden: Leiden University*.

Moulin, D. E., Clark, A. J., Speechley, M. & Morley-Forster, P. K. (2002). Chronic pain in Canada--prevalence, treatment, impact and the role of opioid analgesia. *Pain Research & Management*, *7*(4), 179–184. https://doi.org/10.1155/2002/323085

Myers, S. S., Phillips, R. S., Davis, R. B., Cherkin, D. C., Legedza, A., Kaptchuk, T. J., Hrbek, A., Buring, J. E., Post, D., Connelly, M. T. & Eisenberg, D. M. (2008). Patient expectations as

predictors of outcome in patients with acute low back pain. *Journal of General Internal Medicine*, *23*(2), 148–153. https://doi.org/10.1007/s11606-007-0460-5

Newell, D. & Field, J. (2007). Who will get better? Predicting clinical outcomes in a chiropractic practice. *Clinical Chiropractic*, *10*(4), 179–186. https://doi.org/10.1016/j.clch.2007.10.003

Nielsen, J., Glissmann Nim, C., O'neill, S., Boyle, E., Hartvigsen, J. & Kawchuk, G. N. (2020). Self-reports vs. physical measures of spinal stiffness. *PeerJ*, *8*, 9598. https://doi.org/10.7717/peerj.9598

Niemistö, L., Sarna, S., Lahtinen-Suopanki, T., Lindgren, K. A. & Hurri, H. (2004). Predictive factors for 1-year outcome of chronic low back pain following manipulation, stabilizing exercises, and physician consultation or physician consultation alone. *Journal of Rehabilitation Medicine*, *36*(3), 104–109. https://doi.org/10.1080/16501970310019151

Nim, C. G., Kawchuk, G. N., Schiøttz-Christensen, B. & O'Neill, S. (2020). The effect on clinical outcomes when targeting spinal manipulation at stiffness or pain sensitivity: a randomized trial. *Scientific Reports*, *10*(1), 14615. https://doi.org/10.1038/s41598-020-71557-y

Nordin, M., Carragee, E. J., Hogg-Johnson, S., Weiner, S. S., Hurwitz, E. L., Peloso, P. M., Guzman, J., van der Velde, G., Carroll, L. J., Holm, L. W., Côté, P., Cassidy, J. D. & Haldeman, S. (2009). Assessment of Neck Pain and Its Associated Disorders. Results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. *Journal of Manipulative and Physiological Therapeutics*, *32*(2 SUPPL.), S117–S140. https://doi.org/10.1016/j.jmpt.2008.11.016

Ostelo, R. W. J. G., Deyo, R. A., Stratford, P., Waddell, G., Croft, P., Von Korff, M., Bouter, L. M. & De Vet, H. C. (2008). Interpreting change scores for pain and functional status in low back pain: Towards international consensus regarding minimal important change. *Spine*,

33(1), 90–94. https://doi.org/10.1097/BRS.0b013e31815e3a10

Owens, E. F., DeVocht, J. W., Gudavalli, M. R., Wilder, D. G. & Meeker, W. C. (2007). Comparison of posteroanterior spinal stiffness measures to clinical and demographic findings at baseline in patients enrolled in a clinical study of spinal manipulation for low back pain. *Journal of Manipulative and Physiological Therapeutics*, *30*(7), 493–500. https://doi.org/10.1016/j.jmpt.2007.07.009

Owens Jr, E. F., DeVocht, J. W., Wilder, D. G., Gudavalli, M. R. & Meeker, W. C. (2007). The reliability of a posterior to anterior spinal stiffness measuring system in a population of patients with low back pain. *Journal of Manipulative & Physiological Therapeutics*, *30*(2), 116–123. https://doi.org/10.1016/j.jmpt.2006.12.006

Pagé, I. & Kawchuk, G. (2021). Effects of muscle activity on lumbar spinal stiffness in asymptomatic adults: An investigation using a novel rolling device. *Musculoskeletal Science and Practice*, *52*, 102301. https://doi.org/10.1016/j.msksp.2020.102301

Paris, S. (1985). Physical signs of instability. *Spine*, *10*, 277–279.

Peterson, C. K., Bolton, J. & Humphreys, B. K. (2012). Predictors of improvement in patients with acute and chronic low back pain undergoing chiropractic treatment. *Journal of Manipulative and Physiological Therapeutics*, *35*(7), 525–533. https://doi.org/10.1016/j.jmpt.2012.06.003

Rapoport, J., Jacobs, P., Bell, N. & Klarenbach, S. (2004). Refining the measurement of the economic burden of chronic diseases in Canada. *Chronic Dis Can*, *25*, 13–21.

Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*(2), 147–177. https://doi.org/10.1037/1082-989X.7.2.147

Schopflocher, D. (2003). Chronic pain in Alberta: A portrait from the 1996 National Population Health Survey and the 2001 Canadian Community Health Survey. *Edmonton, Alberta:*

Alberta Health and Wellness.

Schopflocher, D., Taenzer, P. & Jovey, R. (2011). The prevalence of chronic pain in Canada. *Pain Research & Management : The Journal of the Canadian Pain Society = Journal de La Société Canadienne Pour Le Traitement de La Douleur*, *16*(6), 445–450. https://doi.org/10.1155/2011/876306

Seffinger, M. A., Najm, W. I., Mishra, S. I., Adams, A., Dickerson, V. M., Murphy, L. S. & Reinsch, S. (2004). Reliability of Spinal Palpation for Diagnosis of Back and Neck Pain: a systematic review of the literature. *Spine*, *29*(19), E413–E425. https://doi.org/10.1097/01.brs.0000141178.98157.8e

Shirley, D., Ellis, E. & Lee, M. (2002). The response of posteroanterior lumbar stiffness to repeated loading. *Manual Therapy*, *7*(1), 19–25. https://doi.org/10.1054/math.2001.0432

Sions, J. M. & Hicks, G. E. (2017). Back Stiffness Is Associated with Physical Health and Low Back Pain-Related Disability in Community-Dwelling Older Adults. *Pain Medicine (Malden, Mass.)*, *18*(5), 866–870. https://doi.org/10.1093/pm/pnw107

Skargren, E. I. & Öberg, B. E. (1998). Predictive factors for 1-year outcome of low-back and neck pain in patients treated in primary care: Comparison between the treatment strategies chiropractic and physiotherapy. *Pain*, *77*(2), 201–207. https://doi.org/10.1016/S0304-3959(98)00101-8

Snodgrass, S. J., Haskins, R. & Rivett, D. A. (2012). A structured review of spinal stiffness as a kinesiological outcome of manipulation: its measurement and utility in diagnosis, prognosis and treatment decision-making. *J Electromyogr Kinesiol*, *22*(5), 708–723. https://doi.org/10.1186/s12891-019-2543-y

Souza, A. C. de, Alexandre, N. M. C. & Guirardello, E. de B. (2017). Psychometric properties in

instruments evaluation of reliability and validity. *Epidemiologia e Servicos de Saude : Revista Do Sistema Unico de Saude Do Brasil*, *26*(3), 649–659. https://doi.org/10.5123/S1679-49742017000300022

Sran, M. M., Khan, K. M., Zhu, Q. & Oxland, T. R. (2005). Posteroanterior stiffness predicts sagittal plane midthoracic range of motion and three-dimensional flexibility in cadaveric spine segments. *Clinical Biomechanics*, *20*(8), 806–812. https://doi.org/10.1016/j.clinbiomech.2005.04.009

Stanton, T. R. & Kawchuk, G. N. (2009). Reliability of assisted indentation in measuring lumbar spinal stiffness. *Manual Therapy*, *14*(2), 197–205. https://doi.org/10.1016/j.math.2008.01.011

Stanton, T. R., Moseley, G. L., Wong, A. Y. L. & Kawchuk, G. N. (2017). Feeling stiffness in the back: A protective perceptual inference in chronic back pain. *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-09429-1

Stanton, T. R., Latimer, J., Maher, C. G. & Hancock, M. (2009). Definitions of recurrence of an episode of low back pain: A systematic review. *Spine*, *34*(9), 316–322. https://doi.org/10.1097/BRS.0b013e318198d073

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M. & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. *BMJ*, *339*(7713), 157–160. https://doi.org/10.1136/bmj.b2393

Stochkendahl, M. J., Christensen, H. W., Hartvigsen, J., Vach, W., Haas, M., Hestbaek, L., Adams, A. & Bronfort, G. (2006). Manual examination of the spine: a systematic critical literature review of reproducibility. *Journal of Manipulative and Physiological Therapeutics*, *29*(6),

475-485.e10. https://doi.org/10.1016/j.jmpt.2006.06.011

Tamcan, O., Mannion, A. F., Eisenring, C., Horisberger, B., Elfering, A. & Müller, U. (2010). The course of chronic and recurrent low back pain in the general population. *Pain*, *150*(3), 451–457. https://doi.org/10.1016/j.pain.2010.05.019

Thakral, M., Shi, L., Shmerling, R. H., Bean, J. F. & Leveille, S. G. (2014). A stiff price to pay: Does joint stiffness predict disability in an older population? *Journal of the American Geriatrics Society*, *62*(10), 1891–1899. https://doi.org/10.1111/jgs.13070

Tuttle, N., Barrett, R. & Laakso, L. (2008a). Postero-anterior movements of the cervical spine: Repeatability of force displacement curves. *Manual Therapy*, *13*(4), 341–348. https://doi.org/10.1016/j.math.2007.03.005

Tuttle, N., Barrett, R. & Laakso, L. (2008b). Relation between changes in posteroanterior stiffness and active range of movement of the cervical spine following manual therapy treatment. *Spine*, *33*(19), 673–679. https://doi.org/10.1097/BRS.0b013e31817f93f9

Underwood, M. R., Morton, V. & Farrin, A. (2007). Do baseline characteristics predict response to treatment for low back pain? Secondary analysis of the UK BEAM dataset [ISRCTN32683578]. *Rheumatology*, *46*(8), 1297–1302. https://doi.org/10.1093/rheumatology/kem113

Vaillant, M., Pickar, J. G. & Kawchuk, G. N. (2010). Performance and reliability of a variable rate, force/displacement application system. *Journal of Manipulative and Physiological Therapeutics*, *33*(8), 585–593. https://doi.org/10.1016/j.jmpt.2010.08.020

Van Trijffel, E., Anderegg, Q., Bossuyt, P. M. M. & Lucas, C. (2005). Inter-examiner reliability of passive assessment of intervertebral motion in the cervical and lumbar spine: A systematic review. *Manual Therapy*, *10*(4), 256–269. https://doi.org/10.1016/j.math.2005.04.008

Van Tulder, M., Becker, A., Bekkering, T., Breen, A., Del Real, M. T. G., Hutchinson, A., Koes, B., Laerum, E. & Malmivaara, A. (2006). Chapter 3: European guidelines for the management of acute nonspecific low back pain in primary care. *European Spine Journal*, *15*(SUPPL. 2), 169–191. https://doi.org/10.1007/s00586-006-1071-2

Vasseljen, O., Woodhouse, A., Bjørngaard, J. H. & Leivseth, L. (2013). Natural course of acute neck and low back pain in the general population: The HUNT study. *Pain*, *154*(8), 1237–1244. https://doi.org/10.1016/j.pain.2013.03.032

Vavrek, D., Haas, M., Neradilek, M. B. & Polissar, N. (2015). Prediction of pain outcomes in a randomized controlled trial of dose-response of spinal manipulation for the care of chronic low back pain. *BMC Musculoskeletal Disorders*, *16*(1), 1–14. https://doi.org/10.1186/s12891-015-0632-0

Vos, T., Allen, C., Arora, M., Barber, R. M., Brown, A., Carter, A., Casey, D. C., Charlson, F. J., Chen, A. Z., Coggeshall, M., Cornaby, L., Dandona, L., Dicker, D. J., Dilegge, T., Erskine, H. E., Ferrari, A. J., Fitzmaurice, C., Fleming, T., Forouzanfar, M. H., … Zuhlke, L. J. (2016). GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Stud. *The Lancet*, *388*(10053), 1545–1602. https://doi.org/10.1016/S0140-6736(16)31678-6

Waddell, G., Newton, M., Henderson, I., Somerville, D. & Main, C. J. (1993). A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain*, *52*(2), 157–168. https://doi.org/10.1016/0304-3959(93)90127-B

Waddell, G., Somerville, D., Henderson, I. & Newton, M. (1992). Objective Clinical Evaluation of Physical Impairment in Chronic Low Back Pain. *Spine*, *17*(6), 617–628.

https://doi.org/10.1097/00007632-199206000-00001

Walter, S. D., Eliasziw, M. & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Stat Med*, *17*(1), 101–110. https://doi.org/10.1002/(sici)1097-0258(19980115)17:1<101::aid-sim727>3.0.co;2-e

Wells, C., Kolt, G. S., Marshall, P. & Bialocerkowski, A. (2014). The definition and application of pilates exercise to treat people with chronic low back pain: a Delphi survey of Australian physical therapists. *Physical Therapy*, *94*(6), 792–805. https://doi.org/10.2522/ptj.20130030

Werneke, M. & Hart, D. L. (2001). Centralization Phenomenon as a Prognostic Factor for Chronic Low Back Pain and Disability. *Spine*, *26*(7), 758–765. https://doi.org/10.1097/00007632-200104010-00012

Wong, A. Y. L. & Kawchuk, G. N. (2017). The clinical value of assessing lumbar posteroanterior segmental stiffness: a narrative review of manual and instrumented methods. *PM R*, *9*(8), 816–830. https://doi.org/10.1016/j.pmrj.2016.12.001

Wong, A. Y. L., Kawchuk, G., Parent, E. & Prasad, N. (2013a). Within- and between-day reliability of spinal stiffness measurements obtained using a computer controlled mechanical indenter in individuals with and without low back pain. *Manual Therapy*, *18*(5), 395–402. https://doi.org/10.1016/j.math.2013.02.003

Wong, A. Y. L., Kawchuk, G., Parent, E. & Prasad, N. (2013b). Within- and between-day reliability of spinal stiffness measurements obtained using a computer controlled mechanical indenter in individuals with and without low back pain. *Manual Therapy*, *18*(5), 395–402. https://doi.org/10.1016/j.math.2013.02.003

Wong, A. Y. L., Parent, E. C., Prasad, N., Huang, C., Chan, K. M. & Kawchuk, G. N. (2016). Does experimental low back pain change posteroanterior lumbar spinal stiffness and trunk

muscle activity? A randomized crossover study. *Manual Therapy*, *34*, 45–52. https://doi.org/10.1016/j.math.2007.09.013

Wong, A. Y. L., Parent, E. C., Dhillon, S. S., Prasad, N. & Kawchuk, G. N. (2015). Do participants with low back pain who respond to spinal manipulative therapy differ biomechanically from nonresponders, untreated controls or asymptomatic controls? *Spine*, *40*(17), 1329–1337. https://doi.org/10.1097/BRS.0000000000000981

Wong, A. Y. L., Parent, E. C., Dhillon, S. S., Prasad, N., Samartzis, D. & Kawchuk, G. N. (2019). Differential patient responses to spinal manipulative therapy and their relation to spinal degeneration and post-treatment changes in disc diffusion. *European Spine Journal*, *28*(2), 259–269. https://doi.org/10.1007/s00586-018-5851-2

Wu, A., March, L., Zheng, X., Huang, J., Wang, X., Zhao, J., Blyth, F. M., Smith, E., Buchbinder, R. & Hoy, D. (2020). Global low back pain prevalence and years lived with disability from 1990 to 2017: estimates from the Global Burden of Disease Study 2017. *Ann Transl Med*, *8*(6), 299. https://doi.org/10.21037/atm.2020.02.175

Xia, T., Long, C. R., Vining, R. D., Gudavalli, M. R., Devocht, J. W., Kawchuk, G. N., Wilder, D. G. & Goertz, C. M. (2017). Association of lumbar spine stiffness and flexion-relaxation phenomenon with patient-reported outcomes in adults with chronic low back pain – a single-arm clinical trial investigating the effects of thrust spinal manipulation. *BMC Complementary and Alternative Medicine*, *17*(1), 303. https://doi.org/10.1186/s12906-017-1821-1

Young, A., Swain, M. S., Kawchuk, G. N., Wong, A. Y. L. & Downie, A. S. (2020). The bench-top accuracy of the VerteTrack spinal stiffness assessment device. *Chiropr Man Ther*, *28*(1), 42. https://doi.org/10.1186/s12998-020-00331-8

Appendix A. Recruitment Poster

Do you like to lie around?
Are you between 18 and 60 years old?

Volunteers needed to test the reliability of a new instrument for measuring spinal stiffness in asymptomatic participants

Help us develop a new test for back pain research! We are developing **a new instrument to test the stiffness of the spine – a new measure that is becoming** an important indicator of back health. In this study, we need to measure the reliability of the new instrument.

Who can be in the study? You may be eligible to participate in the study if you…

Are between 18 and 60 years of age and have no back pain

Where does the study take place?

Rehabilitation Robotics Lab ECHA- 545, U of Alberta (right by the Health Science LRT)

How much time does it take?

You will lay down on your tummy for ~ 45 minutes on two separate days of the week.
While lying down, the new instrument will measure the stiffness of your back.

Why participate?

Learn about the latest back testing technologies!
Your result may help change the current practice in clinics and research!

Appendix B. Consent and Information Sheet_ Reliability Study

INFORMATION SHEET

Reliability of a new instrument for measuring spinal stiffness in asymptomatic participants

Principal Investigator: Gregory N, Kawchuk, BSc, DC, MSc, PhD

Co-investigators: Maliheh Hadizadeh Bajestani, PT, MSc

Why am I being asked to take part in this research study?

You are being asked to be in this study because you are between 18 and 60 years old and you have no pain in your back or any of the reasons why you cannot be part of this study.

Recently, our lab made a new and cost-effective instrument to measure how stiff your back feels. Our experiment will measure how good this instrument is. This will help other studies to explore how spinal manipulative therapy (SMT) changes back stiffness.

Before you make a decision one of the researchers will go over this form with you. You can always ask questions if you need a better explanation or more information. You will be given a copy of this form to keep.

What is the reason for doing the study?

Low back pain (LBP) is a major cause of pain, disability and cost all over the world. We don't know what causes LBP, but we think it is a mechanical problem.

Recently, our lab has shown that people with LBP have more back stiffness. We have also shown that with treatment, this stiffness can go away in some people but not others. Because of this, we think measuring spinal stiffness will be important for people with back pain. So we can measure spine stiffness in a faster, more accurate way, we have created a new instrument to replace the one we used before. This new instrument (SpineSwiper) costs a lot less to make, can test stiffness in less than 5 minutes and only only takes one person to work the instrument. Now, we need to test the instrument to see that it does all these things.

What will I be asked to do?

This study consists of two parts. Testing of the thoracic spine and testing of the lumbar spine. You are here today for one part only (thoracic or lumbar). You may participate in the other part of the study if you wish and are eligible. You are not required to participate in both parts. Your participation in this study is voluntary. If you decide to participate in this part (thoracic or lumbar), you will visit our lab two times (approximately 1 hour per visit). The two visits should be at the same time of a day and be 1 to 4 days apart. In the first visit, you will fill out some forms and then have an examination of your back. A graduate student researcher who is a trained clinician will perform the examination on your back. We will provide a gown that opens at the back for you to wear during the testing procedures. You can change into the gown in private and keep all your underwear on. If you are wearing a bra, we may ask you to undo it while you are lying on your stomach so we can better test you back. We can give you elastic shorts to wear as well if you like. This gown will let us see and touch the skin of your back from your shoulders to your waistline.

Following the physical examination, you will lie on your tummy and a researcher will touch your spine bones and then mark your skin with an ink pen. This will tell us where we should measure the stiffness of your back. Then, to get you familiar with the instrument, we'll do a practice measurement to see what you think. We'll also do the same in your second visit. During your time at home between your first and second sessions, try to avoid any activities to change the strength of your back muscles so that they have the same strength every time you come to the lab (like sit-ups or core strengthening exercises). You can always do your usual activities – just don't start any new exercises between your visits to our lab.

When we measure your stiffness, a roller wheel will trace a path along your spine in the thoracic section or the lumbar section of you back. (see Figure 1). You will feel a pushing sensation. After we do this, the roller will be removed to let you breath and rest for one minute. We will then do the same thing in the opposite direction. Every time we test your stiffness, we'll ask you you to breathe out then hold your breath that way for about 10 s to relax your back muscles.

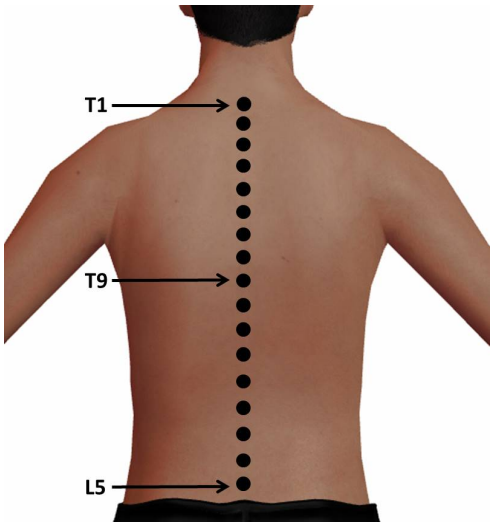


Figure 1. Spine evaluated

After you feel comfortable with the process, we will add a little more weight and you will feel like more pushing against your back. You will always have one minute to rest between stiffness tests. During the test, and afterward, we want you to tell us if there is any pain or discomfort.

You can ask to stop at any time if you feel you cannot continue or do not want to continue. If at any time you want to stop the testing or finish the study, let the investigator know immediately.

What are the risks and discomforts?

You may feel light discomfort in your back following the examination because we will push on your back a bit to see how it feels. You might also feel some discomfort from the roller wheel. Any discomfort you feel should improve on its own within 1-7 days.

You may also experience some side effects that cannot be predicted. Because of this, it is important that you tell one of the researchers listed below right away if you feel any unusual feeling or symptoms or if have any concerns.

What are the benefits to me?

You are not expected to get any benefit or money from being in this research study. We hope that the information we get from doing this study will help us better understand how to better measure spine so that we can use this measure in future tests.

Do I have to take part in the study?

Being in this study is your choice. If you decide to be in the study, you can change your mind and stop being in the study at any time without penalty. In the event of you opt out of the study once we have started collecting data, we will continue to use the data we have collected unless you don't want this to happen. You do not have to answer any questions that you are not comfortable with.

Will I be paid to be in the research?

You will not receive any reimbursement for being in the study.

Will my information be kept private?

Your personal health record related to this study will be kept confidential. Any research data collected about you during this study will not identify you by name, only by a coded number. Your name will not be disclosed outside the research center. Any report published as a result of this study will not identify you by name.

The health information collected in this study will be kept confidential unless release is required by law. All information will be used only for the research study. The researchers and the Health Research Ethics Board may access your study records to monitor the research and verify the accuracy of study information.

In Canada, study information is required to be kept for 7 years. Even if you withdraw from the study, the information and data that is obtained from you for study purposes will not be destroyed.

You have the right to check your health records and request changes if your personal information is incorrect.

What if I have questions?

If you have concerns about your rights as a study participant, you may contact the Research Ethics Office at (780) 492-2615. This office has no affiliation with the study researchers.

Please contact any of the researchers below if you have any questions or concerns:

Gregory Kawchuk	Associate Professor in Physical Therapy
Maliheh Hadizadeh	PhD student in Rehabilitation Medicine

CONSENT FORM

Title of Project: Reliability of a new instrument for measuring spinal stiffness in asymptomatic participants

Principal Investigator(s): Dr. Gregory Kawchuk

Co-investigator(s): Maliheh Hadizadeh Bajestani

_____

Part 2 (to be completed by the research subject):

	Yes	No
Do you understand that you have been asked to be in a research study?	☐	☐
Have you read and received a copy of the attached Information Sheet?	☐	☐
Do you understand the benefits and risks involved in taking part in this research study?	☐	☐
Have you had an opportunity to ask questions and discuss this study?	☐	☐
Do you understand that you are free to withdraw from the study at any time, without having to give a reason and without affecting your future medical care?	☐	☐
Has the issue of confidentiality been explained to you?	☐	☐
Do you understand who will have access to the information you provide?	☐	☐
Do you want the investigator(s) to inform your family doctor that you are participating in this research study? If so, give his/her name _____	☐	☐

Who explained this study to you?_____

I agree to take part in this study: YES ☐ NO ☐

Signature of Research Subject: _____

(Printed Name): _____ Date: _____

I believe that the person signing this form understands what is involved in the study and voluntarily agrees to participate.

Signature of Investigator or Designee: _____ Date _____

THE INFORMATION SHEET MUST BE ATTACHED TO THIS CONSENT FORM AND A COPY GIVEN TO THE RESEARCH SUBJECT

Appendix C. Demographic Information Form

Thank you for completing this questionnaire. This questionnaire will help us to understand your current and past medical condition. Please answer all the questions. There is no right or wrong answer. If you are not sure how to answer a question, you can just give your best answer. Your information will be held in the strictest confidence and can only be accessed by the authorized personnel.

Date…….....

Subject ID…….

Age……....

Gender: □ Female □ Male

Do you have any history of **low back pain**?

□ No □ Yes; If your answer is "Yes", do you know the diagnosis?

……

……

…….....

What is the duration of your current **low back pain**?

……..

Do you have any history of **upper back pain**?

□ No □ Yes; If your answer is "Yes", do you know the diagnosis?

……

……

……

……

What is the duration of your current **upper back pain**?

……

Have you seen any healthcare professionals for current episode?

□ General practitioner □ Physical therapist □ Chiropractor

□ Occupational therapist □ Orthopedic surgeon □ Orthopedic specialist

□ Acupuncturist □ Massage therapist

□ Others…….

Which activity will increase your pain?

□ Sit for..... □ Stand for…….

□ Bend forward for ..... □ Bend forward for……

□ Lay on back for..... □ Lay on stomach for…….

□ Walk on level ground for..... □ Walk uphill for……

□ Cough or sneeze ..... □ Others…….

Which activity will decrease your pain?

□ Sit for..... □ Stand for…….

□ Bend forward for ..... □ Bend forward for…….

□ Lay on back for..... □ Others...…….

Do you have any of these symptoms?

Do you have difficulty in controlling urination?	□ Yes □ No
Do you have difficulty in controlling defecation?	□ Yes □ No
Do you have changes in sensation at your genital region?	□ Yes □ No
Do you have weakness or numbness at your legs?	□ Yes □ No
Do you have unexplained weight loss?	□ Yes □ No
Do you have resting pain that cannot be reduced by painkiller?	□ Yes □ No
Do you have night pain?	□ Yes □ No

Do you have any of the following common health problems that require current medication or medical follow-up? Please specify other unlisted health problems.

Health problem	Have you had this health problem?	Do you need any current medication/treatment?	Medical follow-u
Cancer or tumor	□ Yes □ No	□ Yes □ No	□ Yes □ No
Osteoporosis	□ Yes □ No	□ Yes □ No	□ Yes □ No

Back surgery	□ Yes □ No	□ Yes □ No	□ Yes □ No
Ankylosing spondylitis	□ Yes □ No	□ Yes □ No	□ Yes □ No
Spondylolisthesis	□ Yes □ No	□ Yes □ No	□ Yes □ No
Inflammation of spine	□ Yes □ No	□ Yes □ No	□ Yes □ No
Scoliosis with cobb angle >20°	□ Yes □ No	□ Yes □ No	□ Yes □ No
Congenial spinal disorder	□ Yes □ No	□ Yes □ No	□ Yes □ No
Skin infection/inflammation of back	□ Yes □ No	□ Yes □ No	□ Yes □ No
Pregnancy or suspected pregnancy	□ Yes □ No	□ Yes □ No	□ Yes □ No
Cardiovascular contraindication to exercise	□ Yes □ No	□ Yes □ No	□ Yes □ No
Inflammation of shoulder or arm, which side? □ Left □ Right	□ Yes □ No	□ Yes □ No	□ Yes □ No
Shoulder or arm pain, which side? □ Left □ Right	□ Yes □ No	□ Yes □ No	□ Yes □ No
Arms surgery, which side? □ Left □ Right	□ Yes □ No	□ Yes □ No	□ Yes □ No
Others:	□ Yes □ No	□ Yes □ No	□ Yes □ No
Others:	□ Yes □ No	□ Yes □ No	□ Yes □ No

Social history

Occupation……

Are you participating in competitive sports for more than 3 times per week? □ Yes □ No

Can you lie on stomach for more than 20 minutes? □ Yes □ No

Examiner signature…….

Date……

Appendix D. 11-point Numeric Pain Rating Scale (NRS -11)

Date: .....

Subject ID: .....

No. of visit: □ First session □ Second session

Lower back pain

Please rate your current level of **lower back pain** on the following scale:

□	□	□	□	□	□	□	□	□	□	□
0	1	2	3	4	5	6	7	8	9	10

(no pain) (worst imaginable pain)

Please rate your worst level of **lower back pain** in the last 24 hours on following scale:

□	□	□	□	□	□	□	□	□	□	□
0	1	2	3	4	5	6	7	8	9	10

(no pain) (worst imaginable pain)

Please rate your best level of **lower back pain** in the last 24 hours on following scale:

□	□	□	□	□	□	□	□	□	□	□
0	1	2	3	4	5	6	7	8	9	10

(no pain) (worst imaginable pain)

Upper back pain

Please rate your current level of **upper back pain** on the following scale:

□ □ □ □ □ □ □ □ □ □ □

0 1 2 3 4 5 6 7 8 9 10

(no pain) (worst imaginable pain)

Please rate your worst level of **upper back pain** in the last 24 hours on following scale:

□ □ □ □ □ □ □ □ □ □ □

0 1 2 3 4 5 6 7 8 9 10

(no pain) (worst imaginable pain)

Please rate your best level of **upper back pain** in the last 24 hours on following scale:

□ □ □ □ □ □ □ □ □ □ □

0 1 2 3 4 5 6 7 8 9 10

(no pain) (worst imaginable pain)

During the spinal stiffness test

Please rate your level of **lower back pain** during the spinal stiffness test on the following scale:

□	□	□	□	□	□	□	□	□	□	□
0	1	2	3	4	5	6	7	8	9	10

(no pain) (worst imaginable pain)

Please rate your level of **upper back pain** during the spinal stiffness test on the following scale:

□	□	□	□	□	□	□	□	□	□	□
0	1	2	3	4	5	6	7	8	9	10

(no pain) (worst imaginable pain)

Appendix E. Ethics Approval Form_ Reliability Study

Approval Form

Date:	January 18, 2016	
Study ID:	Pro00061205	
Principal Investigator:	Gregory Kawchuk	
Study Title:	Reliability of a new instrument for measuring spinal stiffness in asymptomatic participants	
Approval Expiry Date:	Tuesday, January 17, 2017	
Approved Consent Form:	Approval Date 1/18/2016	Approved Document Information and Consent Form

Thank you for submitting the above study to the Health Research Ethics Board - Health Panel . Your application, including the following, has been reviewed and approved on behalf of the committee;

- Recruitment Poster v2 (1/16/2016)
- Dempgraphic Data Collection Form (11/28/2015)
- 11 Point Numeric Pain Rating Scale (11/28/2015)
- Body Pain Diagram (11/28/2015)
- Physical Examination Form (11/28/2015)

A renewal report must be submitted next year prior to the expiry of this approval if your study still requires ethics approval. If you do not renew on or before the renewal expiry date, you will have to re-submit an ethics application.

Approval by the Health Research Ethics Board does not encompass authorization to access the patients, staff or resources of Alberta Health Services or other local health care institutions for the purposes of the research. Enquiries regarding Alberta Health Services approvals should be directed to (780) 407-6041. Enquiries regarding Covenant Health should be directed to (780) 735-2274.

Sincerely,

Anthony S. Joyce, Ph.D.
Chair, Health Research Ethics Board - Health Panel

Note: This correspondence includes an electronic signature (validation and approval via an online system).

Appendix F. The First Round of Questions in the Delphi Study

Development of a standard protocol for spinal stiffness measurement using a new loaded rolling wheel system: a Delphi study

Principal Investigator(s) (Supervisor(s)):

Dr. Gregory Kawchuk

Professor

Department of Physical Therapy

Faculty of Rehabilitation Medicine

University of Alberta

Edmonton, AB

Co-investigator(s) (Student(s)):

Maliheh Hadizadeh

Ph.D. Candidate in Rehabilitation Medicine

Faculty of Rehabilitation Medicine

University of Alberta

Edmonton, AB

Invitation to Participate: You are invited to participate in this research study because you have experience with spinal stiffness measurements using VerteTrack. We also assume you have access to e-mail and the Internet over the next 2 months as this study will be performed online.

Purpose of the Study: From this research, we wish to develop an updated best-practice protocol for evaluating spinal stiffness in human participants using VerteTrack to improve the consistency of this assessment in future studies.

Participation: If you wish to participate in this study, please complete the attached survey. The survey should take you approximately 15-20 minutes to complete. You do not have to answer any questions that you do not want to answer. Once you have completed the survey, please choose the "submit" button. We would appreciate receiving it before Sep 14th, 2020. The survey will be closed 2 weeks after the start date. However, we will send everyone notice of reminder one day before closing the survey, regardless of if you've already participated considering the anonymity of the answers.

Benefits: You are not expected to get any benefit from being in this research study. However, some people feel beneficial to share their thoughts. We hope that the information we get from doing this study will help us develop a standard protocol to better measure the spine so that we can use this measure in future tests.

Risks: There are no risks associated with the study. Whether you participate in this study will in no way affect the status that you are entitled to.

Confidentiality and Anonymity: The information that you will share will remain strictly confidential and will be used solely for the purposes of this research. The only people who will have access to the research data are the research team members. Your answers to open-ended questions may be used verbatim in presentations and publications but your name will not be identified. “In order to minimize the risk of security breaches and to help ensure your confidentiality we recommend that you use standard safety measures such as signing out of your account, closing your browser and locking your screen or device when you are no longer using them/ when you have completed the study.” Anonymity is guaranteed since you are not being asked to provide your name or any personal information in the survey.

Data Storage: The survey data files will be securely stored in the PI's password-protected computer with the files encrypted at the faculty of Rehabilitation Medicine, University of Alberta. The data will be kept for 5 years and then destroyed.

Compensation (or Reimbursement): You will not receive any reimbursement for being in the study.

Voluntary Participation: You are under no obligation to participate and if you choose to participate, you may refuse to answer questions that you do not want to answer. Should you choose to withdraw midway through the electronic survey simply close the link and no responses will be included. Your data will be deleted from the database and not used in any analysis if you exit the survey at any time. However, once you submit the survey, the data will be included in the database

and cannot be withdrawn. Given the anonymous nature of the survey, once you have submitted your responses it will no longer be possible to withdraw them from the study.

Contact Information: If you have any questions or require more information about the study itself, please contact Maliheh Hadizadeh, a Ph.D. candidate in Rehabilitation Science, Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB.

The plan for this study has been reviewed by a Research Ethics Board at the University of Alberta (Pro00102734). If you have any questions regarding your rights as a research participant or how the research is being conducted, you may contact the Research Ethics Office at 780-492-2615.

Completion and submission of the survey mean your consent to participate.

Thank you in advance for your time in helping with this study.

I would like to participate in the research project on developing a standard protocol for spinal stiffness measurement using VerteTrack.

☐ Yes ☐ No

Instructions on how to complete Delphi Round 1

Welcome to Delphi round 1.

The first round of this Delphi includes 21 questions in 5 sections:

- Personal experience
- Participants' recruitment
- Participants' safety
- Participants' instructions
- Technical recommendations

All questions are related to the measurement of spinal stiffness in human participants using VerteTrack. You will be presented with some items from the VerteTrack operations manual, and you will be asked to add your answer to the list for each section.

Please provide your free text answers to the following questions. Please give as much detail in your responses as you wish.

Section I_ Personal Experience

1. How many times have you used the VerteTrack device to collect data from participants?
2. How many participants have you tested using the VerteTrack device?
3. Please comment on the ease of use for the VerteTrack?
4. Have you ever had to use the software safety stop or the hardware stop? If so, what were the circumstances?

5. Have you ever had any experiences with VerteTrack equipment malfunction? If so, please explain.
6. Would you use the device again in another study? Why or why not?

Section II_ Participants' Recruitment

7. What inclusion/exclusion criteria might you consider adding to the list below for spinal stiffness measurements using VerteTrack?
    - Pregnancy
    - Serious spinal pathology (e.g., spinal tumor, fracture, infectious disorder, osteoporosis, or other bone demineralizing condition)

Section III_ Participants' Safety

8. What considerations or improvements might you consider adding to the list below to optimize participant safety when data are collected using the Vertetrack device?

    - Make sure all the 1kg weights are removed from the device.
    - Familiarize yourself with the location of the hardware emergency stop (E-stop).
    - Make sure to depress the emergency stop and then disengage it to ensure it is working.

Section IV_ Participants' Instructions

9. What instructions might you consider adding to the list below for participants to follow before undertaking the measurements?
    - You should go to the washroom before testing.
    - You should disrobe/change as necessary to expose the test area sufficiently.
    - You should wear clothes that can be moved to expose your waistline. A gown or shorts might be needed.
    - You have to empty your front and back pockets including coins, keys, cellphones.

10. What instructions might you consider adding to the list below for participants to follow during the measurements?

    - You should not endure discomfort at any time especially when adding weight plates during testing.
    - You should remain still for the duration of the test (~15 min).
    - You should keep arm position the same for the duration of the test.
    - You will be asked to hold your breath at various times during the procedure for an approximate of 10 seconds each time.
    - You should not get up until I tell you.

11. What instructions do you think should be given to the participants after the measurements?
12. If there are multiple measurement sessions in the same and/or different days, what would you recommend participants to do and/or do not do between the sessions?

Section V_ Technical Recommendations

13. What procedures might you consider adding to the list below to familiarize participants with how the VerteTrack collects data?

    - Show participants the orientation video.
    - Practice breathing protocol with the participant before beginning the measurements.
    - Make sure participants have understood the procedure and don't have any questions.

14. What procedure do you recommend for identifying the spinous processes which are used for laser tracking by the device?
15. What procedures do you recommend to ensure the device records from the same starting position?
16. If there are multiple measurement sessions in the same and/or different days, what procedures do you recommend to ensure the participant is oriented in the same position as the previous measurement?
17. What procedures might you consider adding to the list below for placing the wheels over the test area?

    - Move the roller wheels above the highest point of the test area.
    - Make sure there is enough vertical travel in roller to test the most posterior part of the participants' back.
    - Raise the plinth until the highest point on the participant is 3 cm from the wheels.

- Without changing the table height or moving the frame, move the roller wheels to the landing site by positioning the laser over the center of the "X" axis.
- Jog wheel down onto participant and add enough cable slack (approximately 5 extra jogs down).
- Make sure that the wheels are aligned on the skin before running each trial.

18. How do you define a good or a bad trial? What procedures do you apply to ensure a good trial?
19. How would you handle a situation where the software program crashes in the middle of a test for any reason?
20. Is there anything else you would like to tell us about the VerteTrack?
21. Are there any other issues related to the use of the VerteTrack you would you like to see discussed in the next Delphi round?

Now, just a few questions for statistical purposes only.

Present job title: _____

The organization under which the VerteTrack measures were collected: _____

The country in which the VerteTrack measures were performed: _____

Are you...?

- Male
- Female
- Other
- Prefer not to answer

What is your age? _____

What is the highest degree you have completed?

- Bachelor's degree
- Master's degree
- Ph.D.
- Other If other, please specify: _____

How many years of clinical experience do you have in your field? _____

Which health profession do you work in:

- Chiropractic
- Physiotherapy
- Other If other, please specify:_____

Date of survey completion _____

Appendix G. The Second Round of Questions in the Delphi Study

Development of a standard protocol for spinal stiffness measurement using a new loaded rolling wheel system: a Delphi study

In the first round of this Delphi study, we asked about your personal experience with the VerteTrack device and your opinion on participants' recruitment, participants' safety, participants' instructions, and some technical recommendations regarding spinal stiffness measurement using VerteTrack. For the second round, we have provided you with some statements from the first round and we want to know to what extent you agree or disagree with the given statements for each question. We also added four new questions from your recommendations to this round of the study.

The plan for this study has been reviewed by a Research Ethics Board at the University of Alberta (Pro00102734). If you have any questions regarding your rights as a research participant or how the research is being conducted, you may contact the Research Ethics Office at +1 780-492-2615.

Completion and submission of the survey mean your consent to participate.

Thank you in advance for your time in helping with this study.

1. Please indicate to what extent do you agree or disagree with the INCLUSION criteria below for spinal stiffness measurements using VerteTrack.

   (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

   - BMI under 40 for ease of palpation.
   - The ability to tolerate a load of at least 40 N.
   - 18 years or older.
   - Chronic back pain.

Any additional INCLUSION criteria? _____

2. Please indicate to what extent do you agree or disagree with the EXCLUSION criteria below for spinal stiffness measurements using VerteTrack.

   (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

   - Serious spinal pathology (e.g., spinal tumor, fracture, infectious disorder, osteoporosis, or other bone demineralizing condition).
   - Unable to follow instructions (e.g., those with dementia or children (age under 18) who may move during the test.
   - Unable to lie in the prone position (e.g., severe deformities to spine or limbs, static tremor, uncontrolled epilepsy).
   - Unable to maintain their breathing cycle in passive expiration (functional residual capacity) for at least 10 seconds.
   - People with asthma, colds, or breathing disorders.

- A head, neck or thoracoabdominal surgery within the last 6 months.
- History of spine surgery.
- Skin lesion, infection, or open wounds over the back region.
- Previously sacrum trauma/sensitive sacrum.
- Participants with exaggerated spinal curves e.g., thoracic hyper-kyphosis.
- Claustrophobia (a fear of being in closed or small spaces).
- Unstable heart condition.
- Unstable and/or acute disc herniation or injury.
- Scoliosis.
- Spinal canal stenosis.
- Unstable spondylolisthesis.
- Hyperalgesia (an abnormally increased sensitivity to pain).
- Acute pain in the test area.
- Tenderness in the test area.
- Obesity using BMI (e.g., BMI>30).
- Obesity using waist circumference (e.g., waist circumference more than 35 inches in women).
- People who do not feel comfortable with the VerteTrack procedure.
- Pregnancy.

Any additional EXCLUSION criteria? _____

3. What procedures help to FAMILIARIZE PARTICIPANTS with how the VerteTrack collects data?

(1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

- Explain in detail about the duration of the experiment and set of data need to be collected.
- A sensory perception (load on hand).
- Some reassurance that while they may feel pressure on the spine, the device will not cause any harm.
- Explain that there is an emergency stop.
- Allow as many 'practice rounds', with no weights, as the participant needs to feel comfortable.
- Have as many practices runs with no weight as they need to feel comfortable.
- Watch someone else have the measures done (if this is not in the orientation video).
- Show participants the orientation video.
- Practice breathing protocol with the participant before beginning the measurements.
- Make sure participants have understood the procedure and don't have any questions.
- Remind the participants once again some points to note e.g hold breath during the measurement.
- Orientation to the texture and feel of the rolling device.
- Show the device to the participant in person, pointing out the different parts and what their function is to help them further understand the process.

Any additional FAMILIARIZATION procedures? _____

4. Please indicate to what extent do you agree or disagree with providing the following instructions for participants BEFORE ASSESSMENT by the VerteTrack.

(1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

- You should wear comfortable clothing.
- You should wear clothes that can be moved to expose your waistline. A gown or shorts might be needed.
- You should disrobe/change as necessary to expose the test area sufficiently.
- You should remove your glasses.
- Turn your cellphone off.
- You should go to the washroom before testing.
- You have to empty your front and back pockets including coins, keys, cellphones.
- Explain an emergency button can be pushed if needed.
- Explain and practice breathing protocol.
- Explain how to lay down.
- Explain how the device works to increase participant comfort.

Any additional PRE-TEST INSTRUCTIONS? _____

5. To what extent do you agree or disagree with the following procedures for IDENTIFYING THE SPINOUS PROCESSES to be used for laser tracking by the VerteTrack device?

   (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

- Palpation of the spinous processes.
- Check by palpation done by two people.
- Ultrasound.
- Palpation in a prone position in combination with ultrasound for verification.
- Palpation, and confirmation by a healthcare professional.

- Use a standardized palpation procedure based on anatomical landmarks (count up from sacral base and down from T12/ribs) and confirm with diagnostic ultrasound.
- Place hands on iliac crests, identify the L4 spinous process, place a mark on the skin, go down towards the sacrum, identify the L5 spinous process, go up towards the thoracic vertebrae, identify each spinous process, carry out a control on each identified level starting from L5.
- Having someone with sufficient experience landmarking spinous process perform the markings.
- Identify L5 via location 1st sacral tubercle (landing point). Then L5-S1 interspinous up to L1.
- L2 spinous process is at the level of the line joining the inferior borders of the 10th ribs. The intercrestal line is at the level of the L3/4 interspinous space or L3 spinous process.
- It depends on the protocol, the type of study, and the research questions being asked if accurate palpation is needed.

Any additional procedures for IDENTIFYING THE SPINOUS PROCESSES? _____

6. To what extent do you agree or disagree with using the following procedures to ensure the VerteTrack device begins from the same WHEEL STARTING POSITION?"

   (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

- Mark the starting position with a "x".
- Look at the laser from the same angle to ensure it is lined up perfectly before each trial.
- Check the laser goes back to the reference point prior to subsequent runs.

- Make sure the participant is not moving between the trials.
- Photos of the back should be taken.
- Measure the length of the trajectory by a tape measure.

Any additional procedures for the WHEEL STARTING POSITION? _____

7. Please indicate to what extent do you agree or disagree with the statements below regarding PLACING THE WHEELS over the test area.

   (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

- Move the roller wheels above the highest point of the test area.
- Make sure there is enough vertical travel in roller to test the most posterior part of the participants' back.
- Raise the plinth until the highest point on the participant is 3 cm from the wheels.
- Without changing the table height or moving the frame, move the roller wheels to the landing site by positioning the laser over the center of the "X" axis.
- Jog wheel down onto participant and add enough cable slack (approximately 5 extra jogs down).
- Some participants with hyper-lordosis may require more than 5 extra jogs down.
- Make sure that the wheels are aligned on the skin before running each trial.

Any additional recommendations for PLACING THE WHEELS over the test area? _____

8. Please indicate to what extent do you agree or disagree with the statements used DURING ASSESSMENT by the VerteTrack.

(1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

- You should not talk during the procedure.
- You should relax your back and abdominals.
- Let us know if you wish to stop the measurements at any time or if you have any concerns (e.g., discomfort).
- You should give us a sign to indicate that you have exhaled the air and ready to be tested before each trial.
- You should not endure discomfort at any time especially when adding weight plates during testing.
- You should remain still for the duration of the test (~15 min) even when you answer a question in between the trials.
- You should keep arm position the same for the duration of the test.
- You will be asked to hold your breath at various times during the procedure for approximate 10 seconds each time.
- You should wait for my instructions before you move away from the table.

Any additional procedures to follow DURING THE ASSESSMENT? _____

9. To what extent do you agree or disagree that the following POST-TEST INSTRUCTIONS should be given to the participants after assessment by the VerteTrack device?

   (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

- No need for specific instructions after testing. Unless there is interest in the perception of stiffness or mobility in a given study.
- No residue pain or discomfort should remain after the measurements. Any discomfort or problems should be reported to the staff at any time.
- You may experience some mild, short-term pain and discomfort in the area that has been tested.
- It is normal to feel slightly stiff after the measurements.
- You should contact us if you experience any discomfort in the next few hours or days.
- You might feel sore in the next 48hours, this is normal but if the pain does not subside after that time or you feel worried do not hesitate to contact the principal investigator.
- You should walk on a level surface (low-level exercise) for a few minutes after the test procedure.
- Let us know if you feel discomfort after the session or any skin irritation. These two conditions might be expected, but they will eventually disappear.
- Slowly get up and watch your head.
- Wait to get up until the device is removed from above you.
- You may experience some dizziness. If so, sit for a few minutes before standing up.

Any additional POST-TEST INSTRUCTIONS? _____

10. To what extent do you agree or disagree that the following definitions for a GOOD OR A BAD TRIAL when using the VerteTrack device.

    (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

- If the wheels did not move smoothly and they are not continuously pointed forward, it is a bad trial.
- A good trial is a trial where the wheels follow the curvature of the spine without deviating sideways, and which does not cause discomfort to the participant.
- If the displacement decreased at a higher load, it's a bad trial.
- A good trial is one in which the participant is relaxed, does not move, and hold their breath out for the entire trial.
- A bad trial is the one with irregular change in the trajectory line.
- This is typically up to the participant whether the trial is good or bad.
- A good trial is with consistent data collected towards a single participant.
- For a good trial, there is one value for all segments.
- In a good trial, the participant gets an appreciation of how the testing will feel.
- A good or bad trial would be defined based on patient reports and visual inspection.

Any additional definitions for a GOOD OR A BAD TRIAL? _____

11. To what extent do you agree or disagree that the operator should apply the PROCEDURES below to ensure A GOOD TRIAL when using the VerteTrack device?
    (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

- I will monitor the wheels by enough cable slack and will align the wheels.
- I will focus on the graphic trend.
- I will properly communicate with the participant what I expect from them and give them regular feedback.

- I look at the graphics in the software after a few trials to make sure that the graphics look appropriate.
- If I noticed something different with the process, I would mark it as a bad trial.
- I will double-check the data collected before letting the participants leave, repeat if failed.
- I will make sure that the graph output after each trial matches the general graph expected.
- I'll check the values.
- I look for movement, breathing, and tonicity.
- It is necessary that the table on which the patient is positioned has armrests to rest the arms in prone position.

Any additional procedures for A GOOD TRIAL? _____

12. If there are MULTIPLE MEASUREMENT SESSIONS in the same and/or different days, what would you recommend participants to do and/or do not do BETWEEN SESSIONS? (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

- Avoid big meals in between sessions.
- No strenuous exercise should be done in between sessions.
- Go for a walk.
- Use the restroom.
- Sleep well.
- Maintain your normal routine.
- Avoid swimming and scrubbing your back.
- Activities between days depending on the research question.

- Don't undergo any physical demanding activity involving the back.
- You must not have any treatment on the spine between sessions unless this treatment is the subject of experimentation.
- Wear the same clothes for the next session.
- Come back at the same time of the day.
- Don't do any vigorous back exercises two days before the test.
- No additional care between sessions.
- Depending on what is being investigated, might need to control for exercise, food intake, hydration levels (e.g., abdominal contents, gas, delayed onset muscle soreness, etc.).
- Do not do heavy weightlifting/training in between same-day sessions.
- Recommendations to be more or less active than usual could be a confounding factor to results.
- Do not begin new physically intensive activities between measurement sessions.
- If you take medication like muscle relaxants or pain killers, take the medication after the assessment.

Any additional BETWEEN SESSIONS recommendations for participants? _____

13. If there are MULTIPLE MEASUREMENT SESSIONS in the same and/or different days, what procedures do you recommend to ensure the participant is oriented in the SAME POSITION as the previous measurement?

    (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

- Participants should feel just as comfortable as before.

- Put a band aid/ adhesive tape on top of the marked "x" spot so you don't lose it for the next visit.
- Take a photo with the consent of the participant.
- Measure the trajectory distance.
- Tape on the table and on the floor to ensure the same position of equipment and person on the table.
- Since the testing plinth has a hole, the participant will always align at approximately the same distance from the cephalic end of the plinth.
- Use a permanent marker (particularly for S1) to ensure the starting position of measurement is the same.
- Have a standardized examination table with markings that could be used to align participants in a reproducible manner.
- Keep the reference points intact.
- Take notes on the position of the patient (head, arms, legs).

Any additional procedures for orienting the participant in the SAME POSITION in MULTIPLE MEASUREMENT SESSIONS? _____

14. How would you handle a situation where the SOFTWARE PROGRAM CRASHES in the middle of a test for any reason?

    (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

- I will stop the software and restart the software.
- I will close the software and restart the computer.

- I will turn off the control box and restart the whole system.
- I will remove the device from above participant and start over.
- I will re-calibrate the device.
- I will re-schedule the participant.
- I will inform the participant from the situation and will ask to lie still for the issue to be fixed.
- I will ask participant's permission to start over.
- I will ask participants if they would like a rest before starting over.
- I will re-do the problematic trial and resume the measurements.
- I will re-do the measurements from 0N.
- I will press the emergency stop button.
- I will remove all the weights.
- Make sure the participant is safely out of the device.

Any additional recommendation on how to handle the SOFTWARE PROGRAM CRASHES?

_____

15. Please indicate to what extent do you agree or disagree with the statements below regarding optimizing PARTICIPANT SAFETY when data are collected using the VerteTrack device. (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

- The safety stop button should immediately elevate the load and return the rolling arm to a position away from the patient - so that the patient can exit if needed.
- Clear instructions to participants with expectations explained.

- The participants should not get up before the frame is off them.
- Make sure to remove the weights one by one at the end of the measurement.
- Make sure the device is properly operational (or locked in place) when loading weights.
- Make sure all the 1kg weights are removed from the device before and after assessment by the VerteTrack.
- Familiarize yourself with the location of the hardware emergency stop (E-stop) before assessment by the VerteTrack.
- Make sure to depress the emergency stop and then disengage it to ensure it is working before assessment by the VerteTrack.
- Have a mirror to be able to see the client's face.
- Have an easy reading format for clients with disabilities before assessment by the VerteTrack.
- Follow the suggested pre-test protocol to make sure all "detectors" are functioning properly.
- Procedures explained to participants for emergency stop.
- Continuing to check in with the patient throughout the process to make sure that they are feeling okay.
- Disinfect the wheels/bench/equipment prior to each participant.

  Any recommendations for optimizing PARTICIPANT SAFETY? _____

Please provide your free text answers to the four following questions. Please give as much detail in your responses as you wish.

16. With respect to pregnancy, how details you think the exclusion criteria should be (e.g., from the first day of pregnancy to the day before delivery, from the first day of pregnancy till 6 weeks following delivery, only last trimester, etc.)?

17. Do you have any recommendations for a DEMONSTRATION VIDEO on the use of VerteTrack?

18. In the event of a software program crash, what would you like to see in a TROUBLESHOOTING GUIDE?

19. What would you like to see for the FUTURE STUDIES using the VerteTrack?

Appendix H. The Third Round of Questions in the Delphi Study

Development of a standard protocol for spinal stiffness measurement using a new loaded rolling wheel system: a Delphi study

Thank you once again for completing the second round of our Delphi study. Now in this third and final round, we present you with the agreement results to date. So far 115/171 items have achieved our desired level of 70% agreement. Statements having 70% or greater level of agreement will be retained in the recommended VerteTrack Protocol. While these items have reached agreement, you may re-rate them if you like. You will also see statements having less than 70% agreement. These items must be re-rated in this final round of the

Delphi study. If consensus is not reached, they will be dropped from the recommended VerteTrack Protocol.

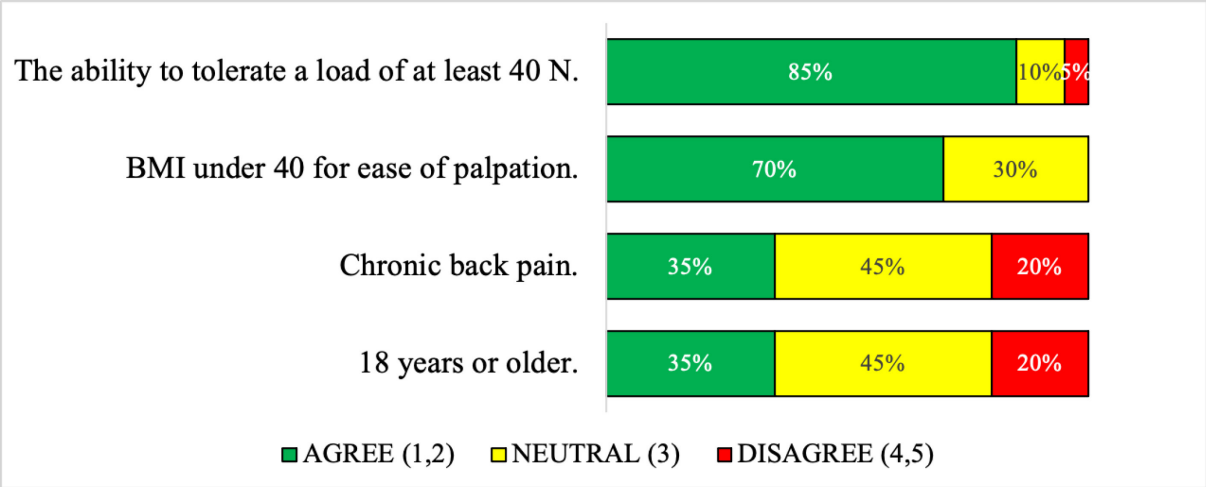
The plan for this study has been reviewed by a Research Ethics Board at the University of Alberta (Pro00102734). If you have any questions regarding your rights as a research participant or how the research is being conducted, you may contact the Research Ethics Office at +1 780-492-2615.

Completion and submission of the survey mean your consent to participate.

Thank you in advance for your time in helping with this study.

1. Please indicate to what extent do you agree or disagree with the INCLUSION criteria below for spinal stiffness measurements using VerteTrack.

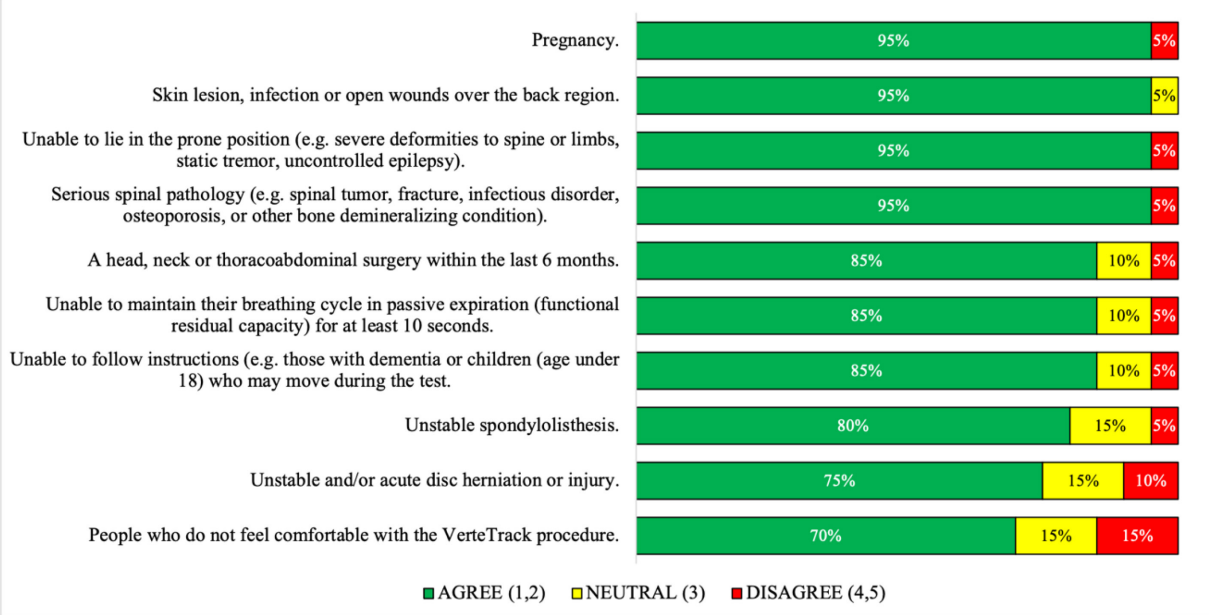
   (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)



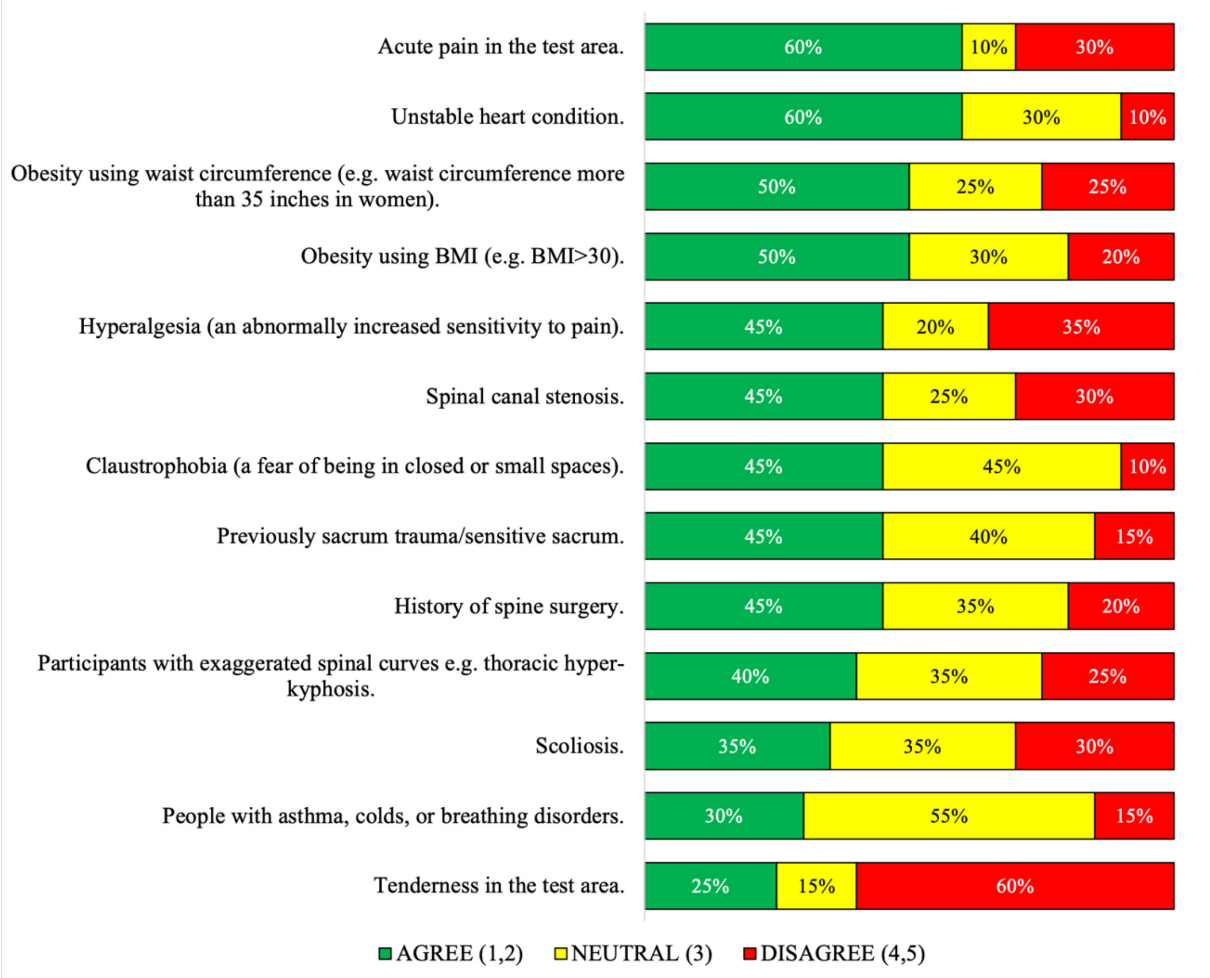
- The ability to tolerate a load of at least 40 N (85% agreement).
- BMI under 40 for ease of palpation (70% agreement).
- Chronic back pain (35% agreement).
- 18 years or older (35% agreement).

2. Please indicate to what extent do you agree or disagree with the EXCLUSION criteria below for spinal stiffness measurements using VerteTrack.

   (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)



- Pregnancy (95% agreement).
- Skin lesion, infection, or open wounds over the back region (95% agreement).
- Unable to lie in the prone position (e.g., severe deformities to spine or limbs, static tremor, uncontrolled epilepsy) (95% agreement).
- Serious spinal pathology (e.g., spinal tumor, fracture, infectious disorder, osteoporosis, or other bone demineralizing condition) (95% agreement).
- A head, neck or thoracoabdominal surgery within the last 6 months (85% agreement).
- Unable to maintain their breathing cycle in passive expiration (functional residual capacity) for at least 10 seconds (85% agreement).
- Unable to follow instructions (e.g., those with dementia or children (age under 18) who may move during the test (85% agreement).
- Unstable spondylolisthesis (80% agreement).
- Unstable and/or acute disc herniation or injury (75% agreement).
- People who do not feel comfortable with the VerteTrack procedure (70% agreement).



- Acute pain in the test area (depends on whether a participant can tolerate the loading and how long the aggravated pain will subside) (60% agreement).
- Unstable heart condition (60% agreement).
- Obesity using waist circumference (e.g., waist circumference more than 35 inches in women) (50% agreement).
- Obesity using BMI (e.g., BMI>30) (50% agreement).
- Hyperalgesia (an abnormally increased sensitivity to pain) (45% agreement).
- Spinal canal stenosis (45% agreement).

- Claustrophobia (a fear of being in closed or small spaces) (45% agreement).
- Previously sacrum trauma/sensitive sacrum (45% agreement).
- History of spine surgery (depends on whether a participant can tolerate the loading and how long the aggravated pain will subside) (45% agreement).
- Participants with exaggerated spinal curves e.g., thoracic hyper-kyphosis (40% agreement).
- Scoliosis (35% agreement).
- People with asthma, colds, or breathing disorders (30% agreement).
- Tenderness in the test area (depends on whether a participant can tolerate the loading and how long the aggravated pain will subside) (25% agreement).

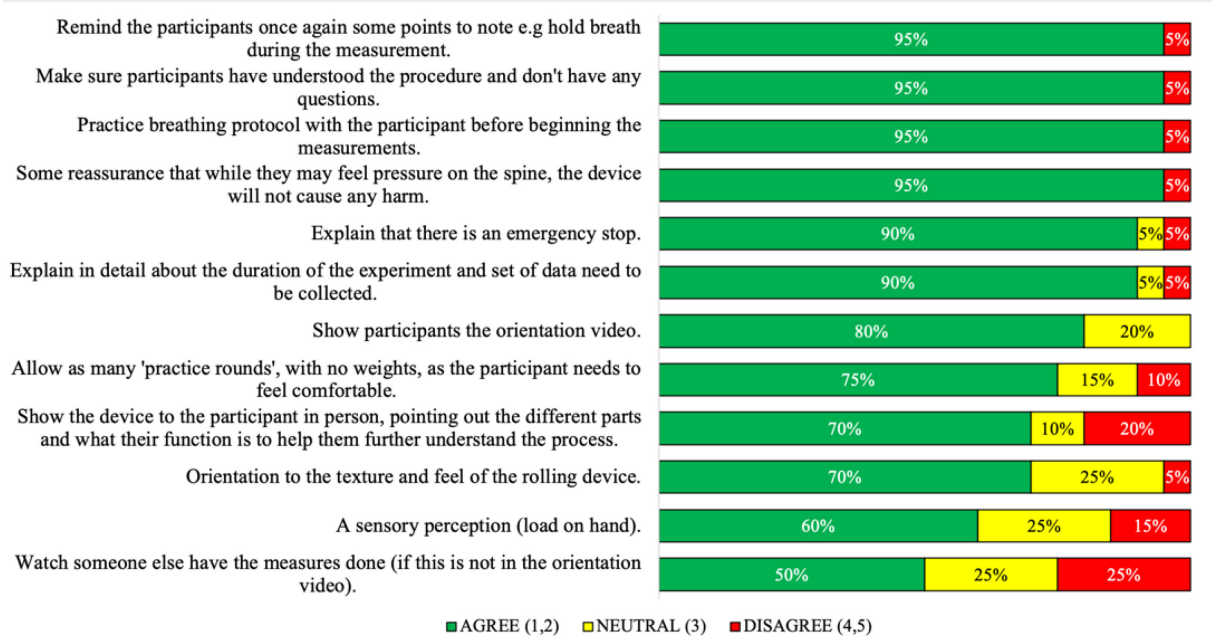
3. Please rate each of the following new items for their inclusion as more detailed statements regarding pregnancy.

   (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

- Pregnant woman should not participate at any stage of pregnancy
- From the first day of pregnancy to the day following the delivery.
- From the first day of pregnancy till 1 month postpartum.
- From confirmation of pregnancy till 6 weeks postpartum.
- From the first day of pregnancy to 3 months postpartum.
- From confirmation of pregnancy to 12 months postpartum.
- Excluded from the second trimester.

4. What procedures help to FAMILIARIZE PARTICIPANTS with how the VerteTrack collects data?

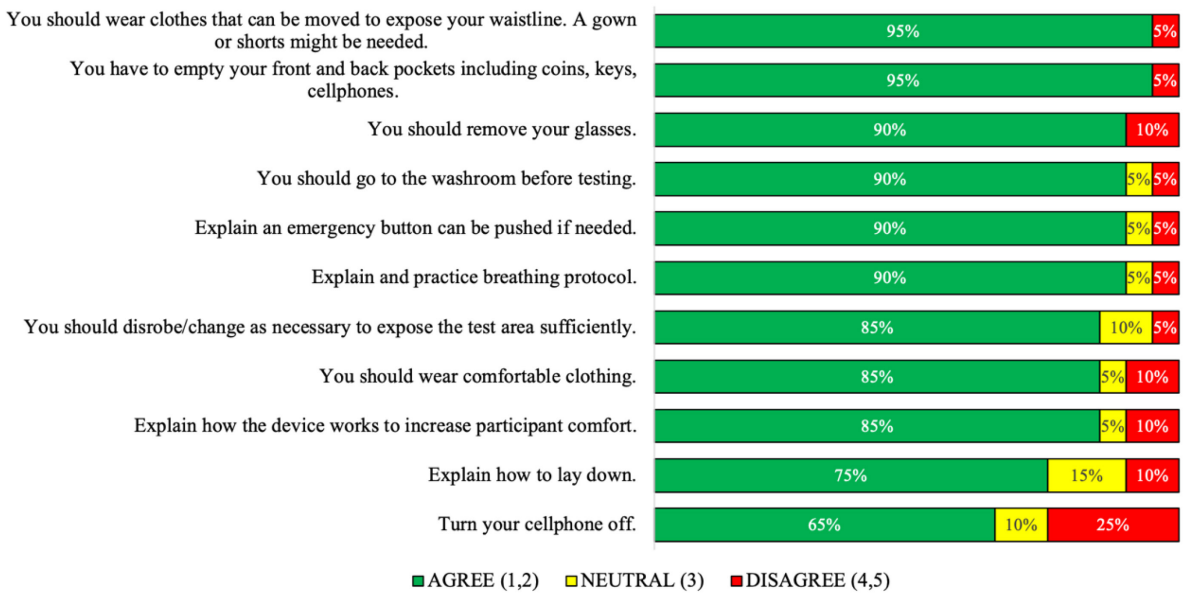
   (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)



- Remind the participants once again some points to note e.g., hold breath during the measurement (95% agreement).
- Make sure participants have understood the procedure and don't have any questions (95% agreement).
- Practice breathing protocol with the participant before beginning the measurements (95% agreement).
- Some reassurance that while they may feel pressure on the spine, the device will not cause any harm (95% agreement).

- Explain that there is an emergency stop (90% agreement).
- Explain in detail about the duration of the experiment and set of data need to be collected (90% agreement).
- Show participants the orientation video (80% agreement).
- Allow an upper limit of 5 unloaded practice rounds and always note in the protocol how many practice rounds were completed (75% agreement).
- Show the device to the participant in person, pointing out the different parts and what their function is to help them further understand the process (70% agreement).
- Orientation to the texture and feel of the rolling device (70% agreement).
- A sensory perception (load on hand) (60% agreement).
- Watch someone else have the measures done (if this is not in the orientation video) (50% agreement).

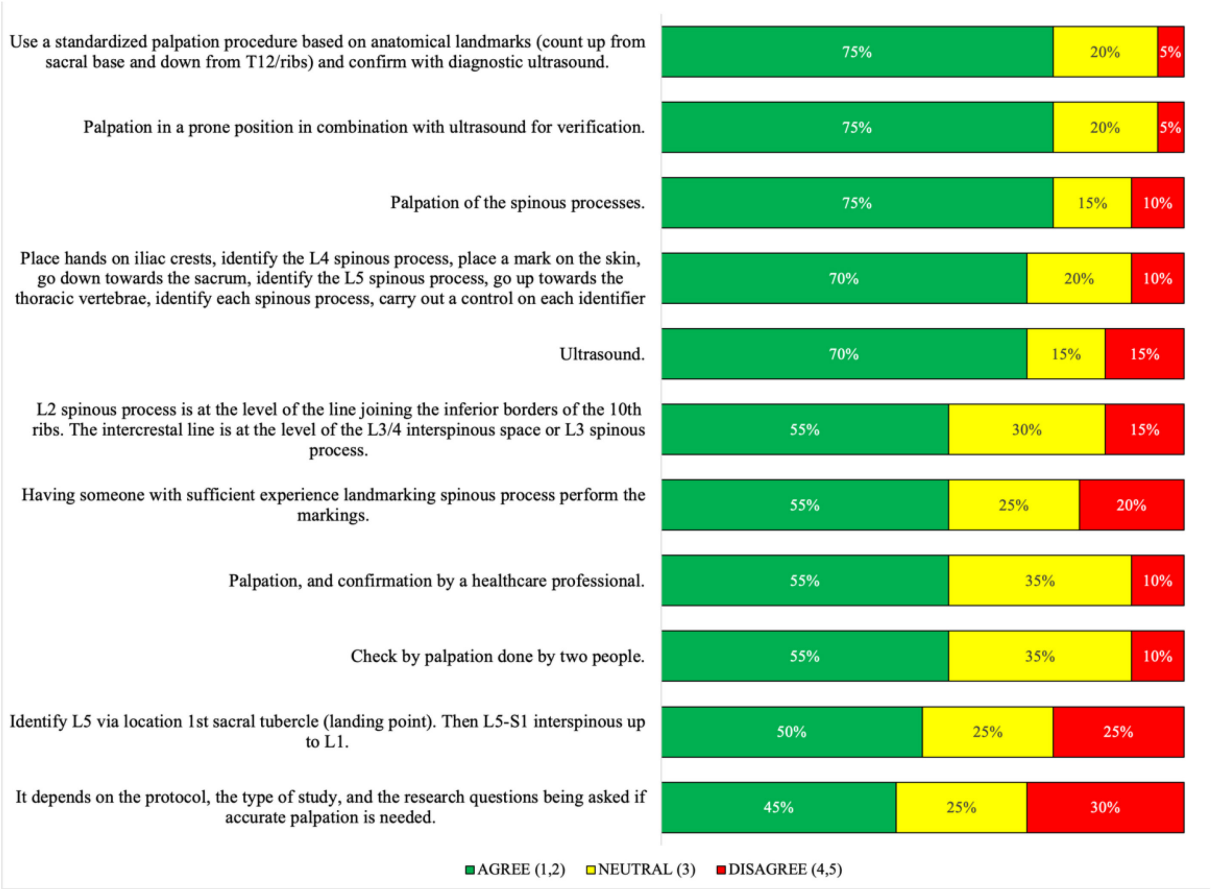
5. Please indicate to what extent do you agree or disagree with providing the following instructions for participants BEFORE ASSESSMENT by the VerteTrack.
   (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)



- You should wear clothes that can be moved to expose your waistline. A gown or shorts might be needed (95% agreement).
- You have to empty your front and back pockets including coins, keys, cellphones (95% agreement).
- You should remove your glasses (90% agreement).
- You should go to the washroom before testing (90% agreement).
- Explain some circumstances where the participant might want to press the emergency stop. E.g., if they have a radicular pain, and they experience pain into their leg (90% agreement).
- Explain and practice breathing protocol (90% agreement).
- You should disrobe/change as necessary to expose the test area sufficiently (85% agreement).
- You should wear comfortable clothing (85% agreement).
- Explain how the device works to increase participant comfort (85% agreement).
- Explain how to lay down (75% agreement).

- Cell phones should be allowed to stay on for emergency calls etc. but the participant should be instructed that we don't want them looking at their phones during the protocol (65% agreement).

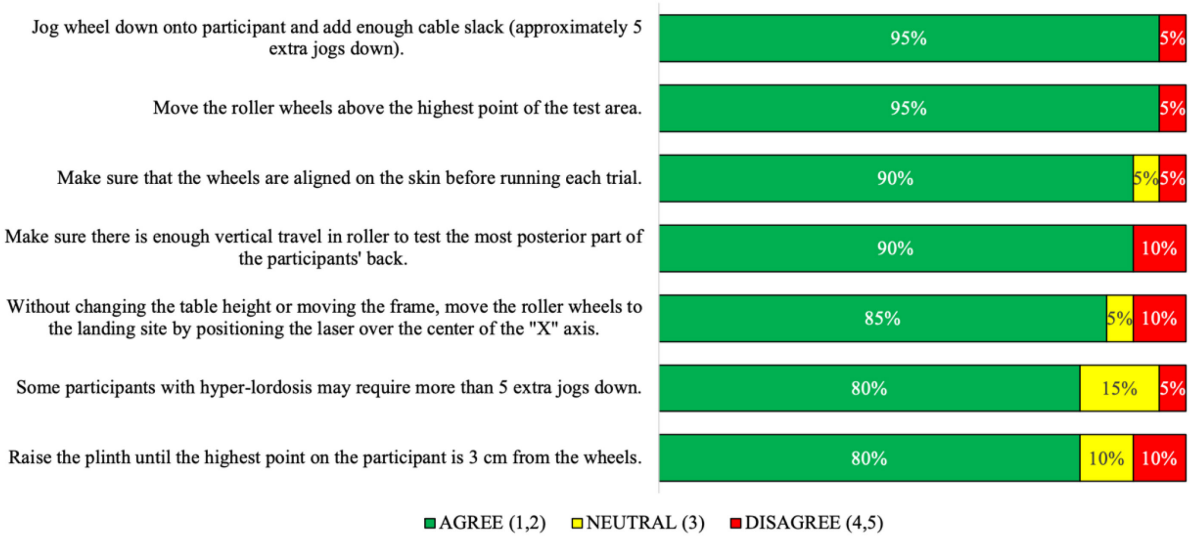
6. To what extent do you agree or disagree with the following procedures for IDENTIFYING THE SPINOUS PROCESSES to be used for laser tracking by the VerteTrack device? (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)



- Use a standardized palpation procedure based on anatomical landmarks (count up from sacral base and down from T12/ribs) and confirm with diagnostic ultrasound (75% agreement).
- Palpation in a prone position in combination with ultrasound for verification (75% agreement).
- Palpation of the spinous processes (75% agreement).
- Place hands on iliac crests, identify the L4 spinous process, place a mark on the skin, go down towards the sacrum, identify the L5 spinous process, go up towards the thoracic vertebrae, identify each spinous process (70% agreement).
- Ultrasound if available (70% agreement).
- L2 spinous process is at the level of the line joining the inferior borders of the 10th ribs. The intercristal line is at the level of the L3/4 interspinous space or L3 spinous process (55% agreement).
- Having someone with sufficient experience landmarking spinous process perform the markings (55% agreement).
- Palpation, and confirmation by a healthcare professional (55% agreement).
- Check by palpation done by two people (55% agreement).
- Identify L5 via location 1st sacral tubercle (landing point). Then L5-S1 interspinous up to L1(50% agreement).
- It depends on the protocol, the type of study, and the research questions being asked if accurate palpation is needed (45% agreement).

7. Please indicate to what extent do you agree or disagree with the statements below regarding PLACING THE WHEELS over the test area.

   (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

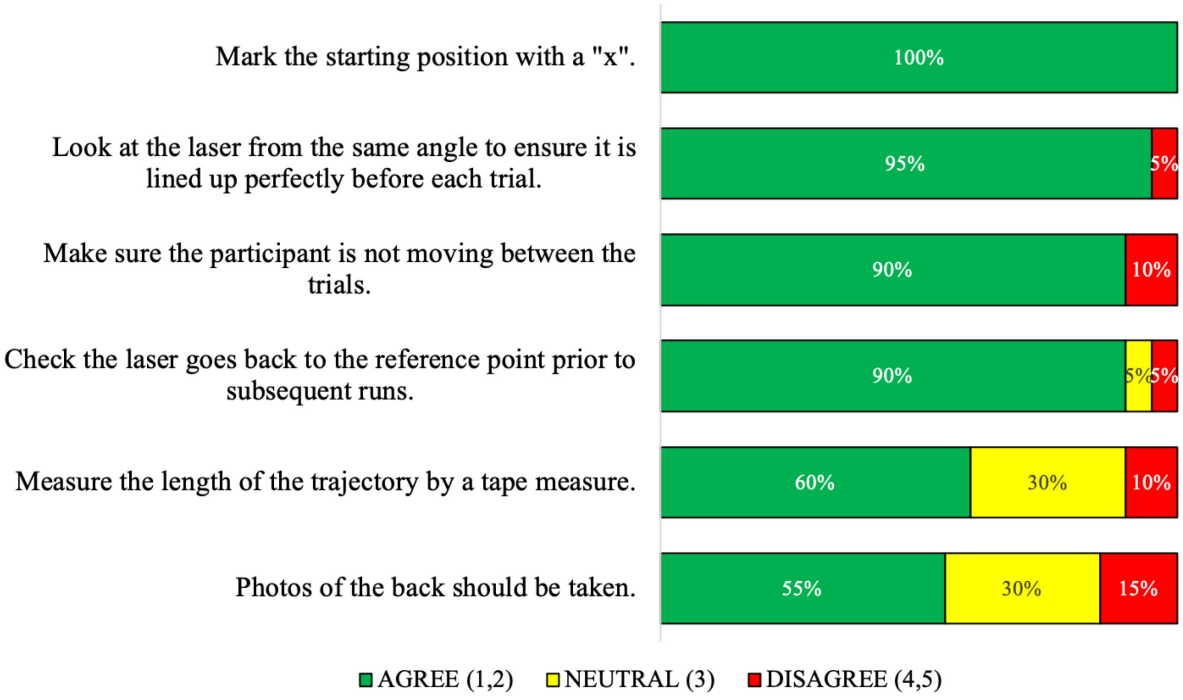


- Jog wheel down onto participant and add enough cable slack (approximately 5 extra jogs down) (95% agreement).
- Move the roller wheels above the highest point of the test area (95% agreement).
- Make sure that the wheels are aligned on the skin before running each trial (90% agreement).
- Make sure there is enough vertical travel in roller to test the most posterior part of the participants' back (90% agreement).
- Without changing the table height or moving the frame, move the roller wheels to the landing site by positioning the laser over the center of the "X" axis (85% agreement).
- Some participants with hyper-lordosis may require more than 5 extra jogs down (80% agreement).

- Raise the plinth until the highest point on the participant is 3 cm from the wheels (80% agreement).
- Identify the slack in the cable before wheels come up from skin, make sure the wheels go deep enough into the skin, and look for slight inhales or slight movements by participant (new, no bar graph).

8. To what extent do you agree or disagree with using the following procedures to ensure the VerteTrack device begins from the same WHEEL STARTING POSITION?"

   (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

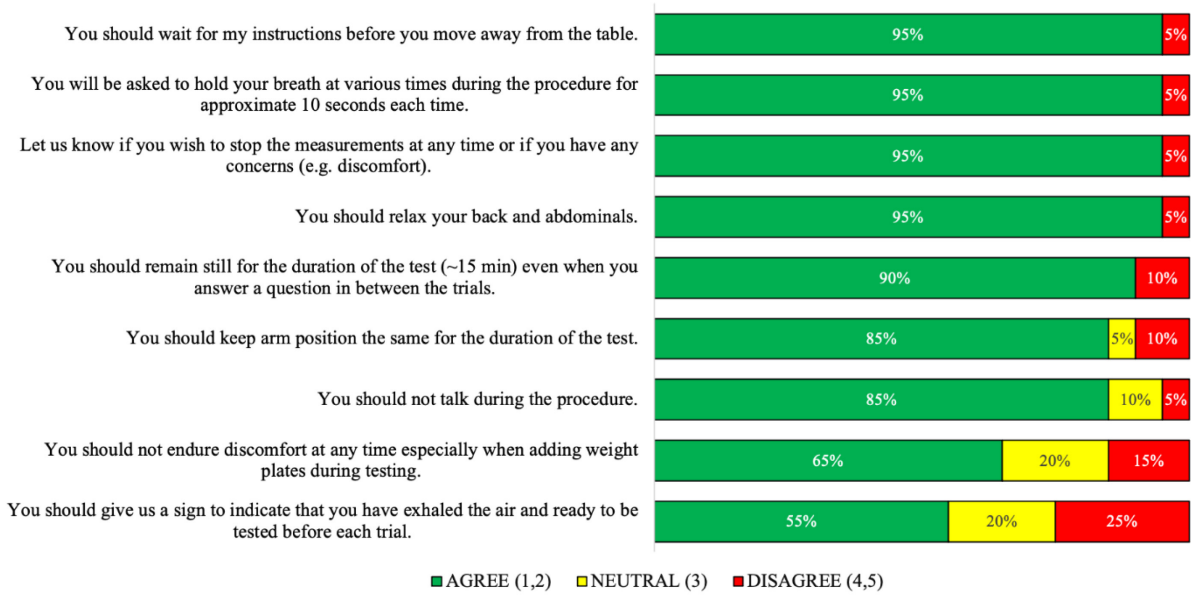


- Mark the starting position with a "x" (100% agreement).

- Look at the laser from the same angle to ensure it is lined up perfectly before each trial (95% agreement).
- Make sure the participant is not moving between the trials (90% agreement).
- Check the laser goes back to the reference point prior to subsequent runs (90% agreement).
- Measure the length of the trajectory by a tape measure (60% agreement).
- Photos of the back should be taken (55% agreement).

9. Please indicate to what extent do you agree or disagree with the statements used DURING ASSESSMENT by the VerteTrack.

   (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

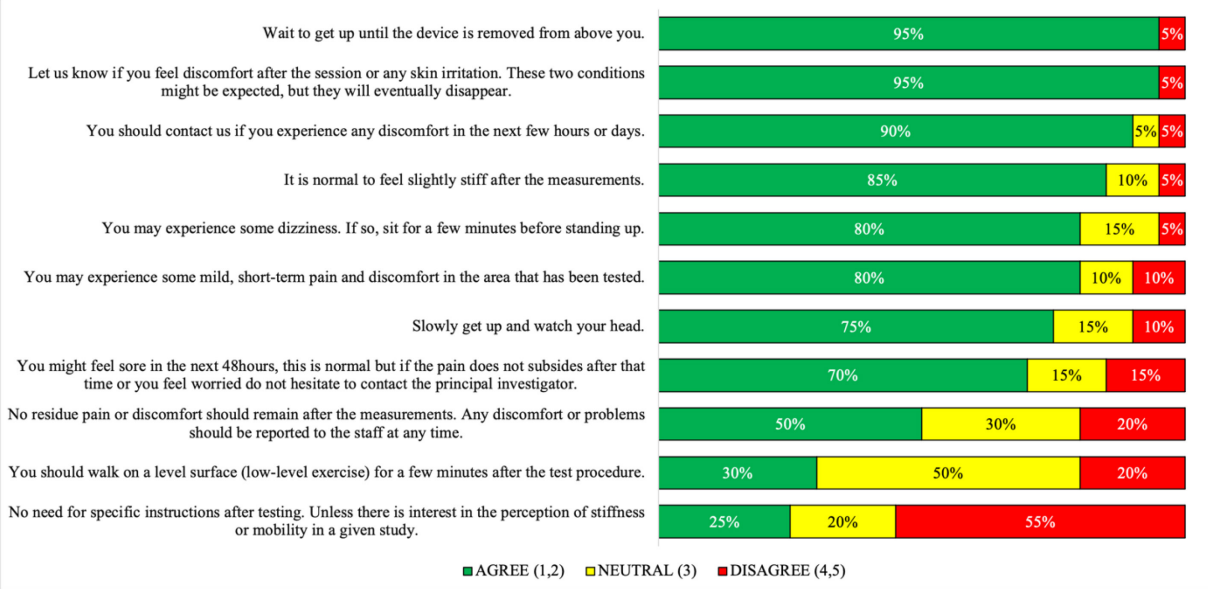


- You should wait for my instructions before you move away from the table (95% agreement).

- You will be asked to hold your breath at various times during the procedure for approximate 10 seconds each time (95% agreement).
- Let us know if you wish to stop the measurements at any time or if you have any concerns (e.g., discomfort) (95% agreement).
- You should relax your back and abdominals (95% agreement).
- You should remain still for the duration of the test (~15 min) even when you answer a question in between the trials (90% agreement).
- You should keep arm position the same for the duration of the test (85% agreement).
- You should not talk during the procedure (85% agreement).
- You should not endure discomfort at any time especially when adding weight plates during testing (65% agreement).
- You should give us a sign to indicate that you have exhaled the air and ready to be tested before each trial (55% agreement).
- You should use the emergency stop button if you cannot tolerate the testing procedure (new, no bar graph).
- The operator should check the participant's readiness for each trial (new, no bar graph).
- You'll be instructed when you can start breathing again (new, no bar graph).

10. To what extent do you agree or disagree that the following POST-TEST INSTRUCTIONS should be given to the participants after assessment by the VerteTrack device?

    (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

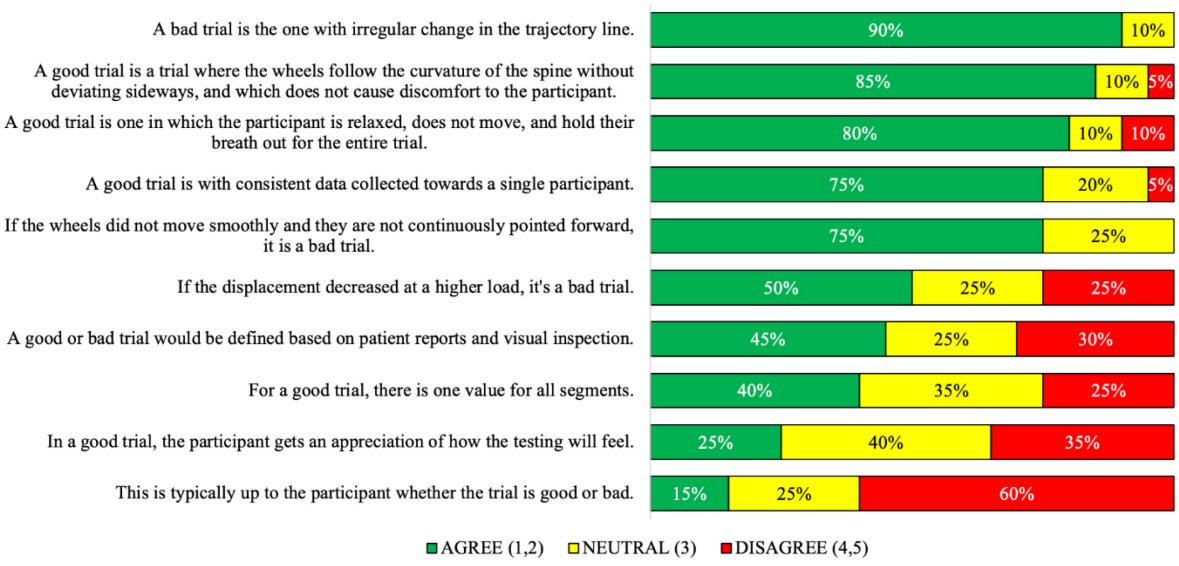


- Wait to get up until the device is removed from above you (95% agreement).
- Let us know if you feel discomfort after the session or any skin irritation. These two conditions might be expected, but they will eventually disappear (95% agreement).
- You should contact us if you experience any discomfort in the next few hours or days (90% agreement).
- It is normal to feel slightly stiff after the measurements (85% agreement).
- You may experience some dizziness. If so, sit for a few minutes before standing up (80% agreement).
- You may experience some mild, short-term pain and discomfort in the area that has been tested (80% agreement).
- Slowly get up and watch your head (75% agreement).
- You might feel sore in the next 48hours, this is normal but if the pain does not subside after that time or you feel worried do not hesitate to contact the principal investigator (70% agreement).

- No residue pain or discomfort should remain after the measurements. Any discomfort or problems should be reported to the staff at any time (50% agreement).
- You should walk on a level surface (low-level exercise) for a few minutes after the test procedure (30% agreement).
- No need for specific instructions after testing. Unless there is interest in the perception of stiffness or mobility in a given study (25% agreement).

11. To what extent do you agree or disagree that the following definitions for a GOOD OR A BAD TRIAL when using the VerteTrack device.

    (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

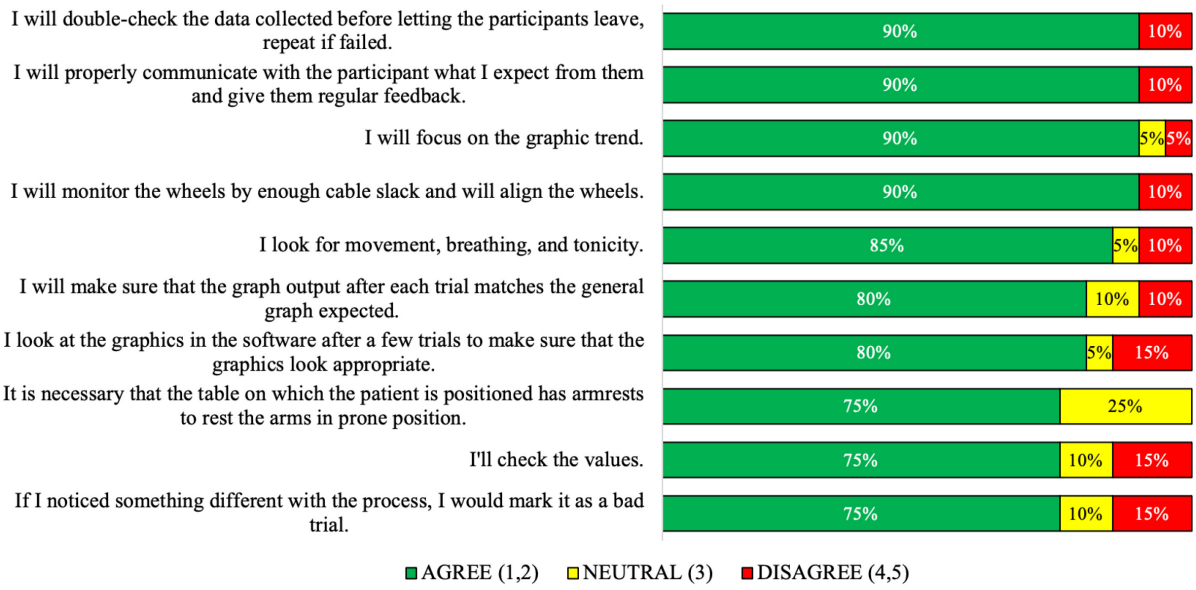


- A bad trial is the one with irregular change in the trajectory line (90% agreement).

- A good trial is a trial where the wheels follow the curvature of the spine without deviating sideways, and which does not cause discomfort to the participant (85% agreement).
- A good trial is one in which the participant is relaxed, does not move, and hold their breath out for the entire trial (80% agreement).
- A good trial is when consistent data is collected for each participant (75% agreement).
- If the wheels did not move smoothly and they are not continuously pointed forward, it is a bad trial (75% agreement).
- If the displacement decreased at a higher load, it's a bad trial (50% agreement).
- A good or bad trial would be defined based on patient reports and visual inspection (45% agreement).
- A good trial is when same value is collected for all segments (40% agreement).
- In a good trial, the participant gets an appreciation of how the testing will feel (25% agreement).
- This is typically up to the participant whether the trial is good or bad (15% agreement).

12. To what extent do you agree or disagree that the operator should apply the PROCEDURES below to ensure A GOOD TRIAL when using the VerteTrack device?

    (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

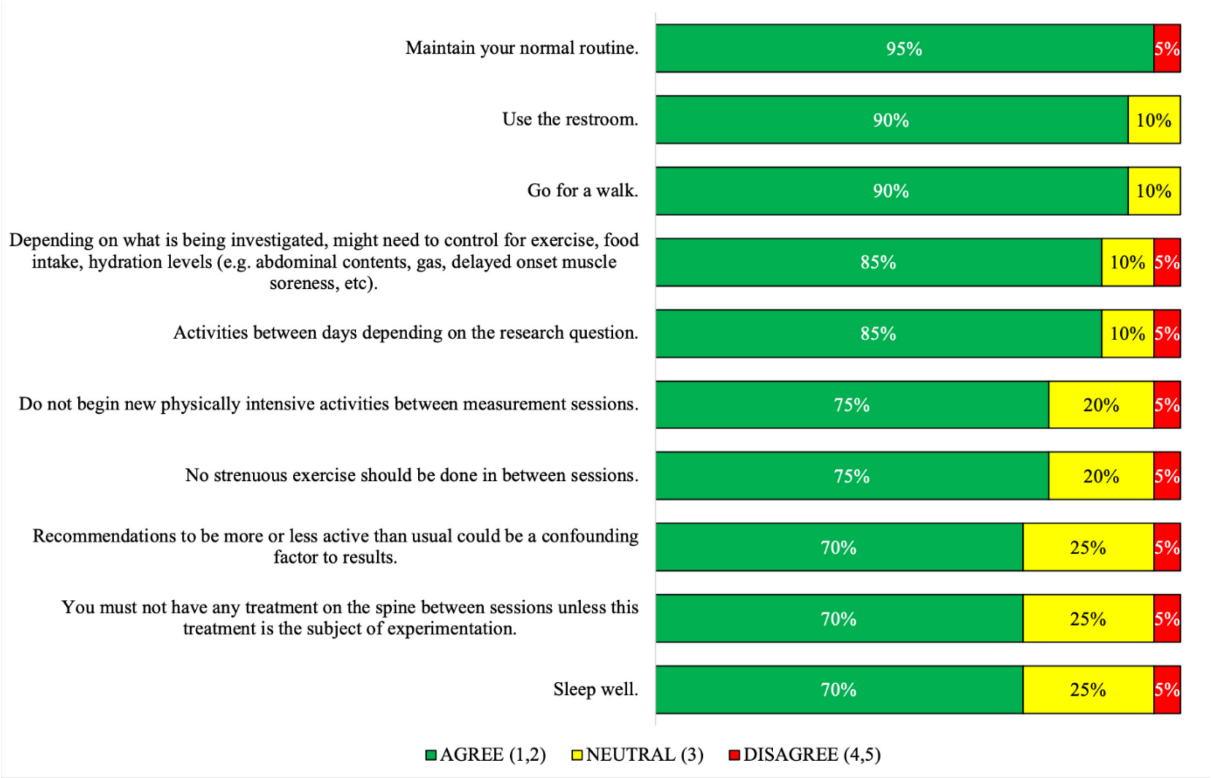


- I will double-check the data collected before letting the participants leave, repeat if failed (90% agreement).
- I will properly communicate with the participant what I expect from them and give them regular feedback (90% agreement).
- I will focus on the graphic trend (90% agreement).
- I will monitor the wheels by enough cable slack and will align the wheels (90% agreement).
- I look for movement, breathing, and tonicity (85% agreement).
- I will make sure that the graph output after each trial matches the general graph expected (80% agreement).
- I look at the graphics in the software after a few trials to make sure that the graphics look appropriate (80% agreement).
- It is necessary that the table on which the patient is positioned has armrests to rest the arms in prone position (75% agreement).
- I'll check the values (75% agreement).

- If I noticed something different with the process, I would mark it as a bad trial (75% agreement).

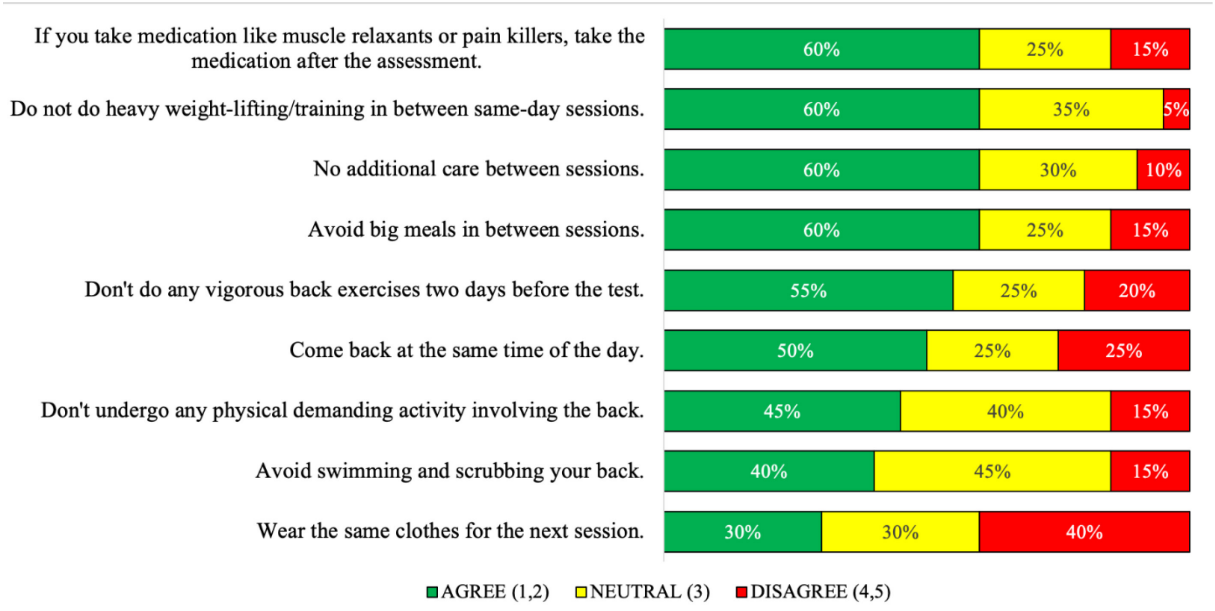
13. If there are MULTIPLE MEASUREMENT SESSIONS in the same and/or different days, what would you recommend participants to do and/or do not do BETWEEN SESSIONS?

(1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)



- Maintain your normal routine (95% agreement).
- Use the restroom (90% agreement).
- Go for a walk (90% agreement).

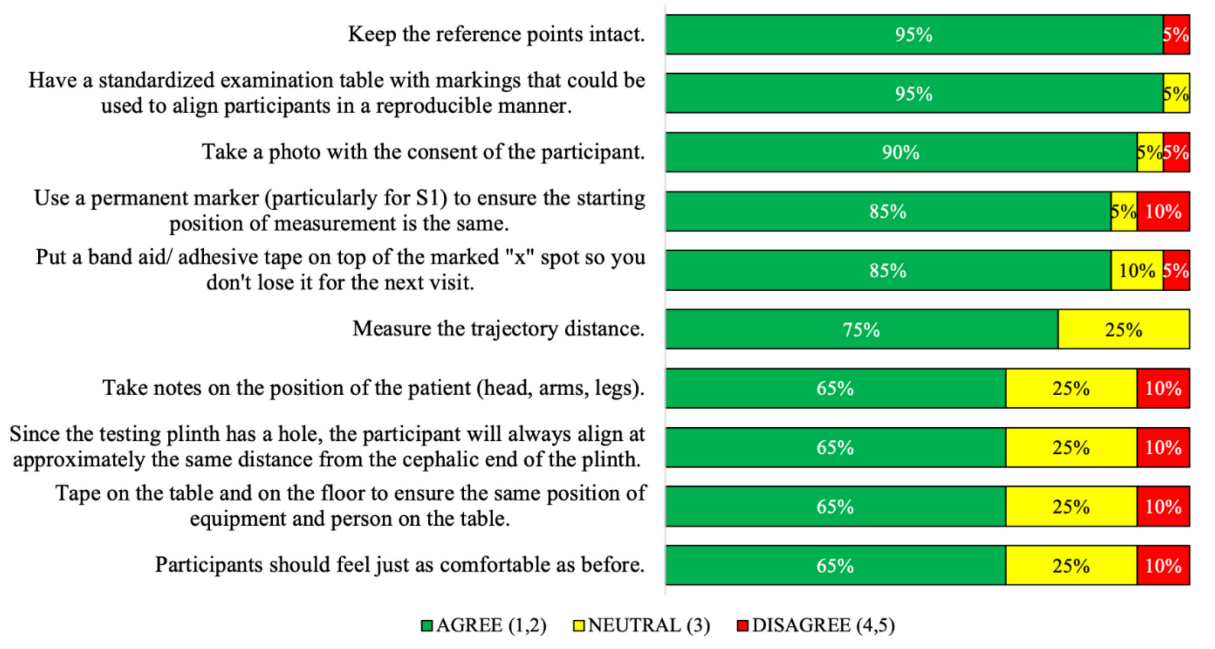
- Depending on what is being investigated, might need to control for exercise, food intake, hydration levels (e.g. abdominal contents, gas, delayed onset muscle soreness, etc) (85% agreement).
- Activities between days depending on the research question (85% agreement).
- Do not begin new physically intensive activities between measurement sessions (75% agreement).
- No strenuous exercise should be done in between sessions (75% agreement).
- Recommendations to be more or less active than usual could be a confounding factor to results (70% agreement).
- You must not have any treatment on the spine between sessions unless this treatment is the subject of experimentation (70% agreement).
- Sleep well (70% agreement).



- If you take medication like muscle relaxants or pain killers, take the medication after the assessment (60% agreement).
- Do not do heavy weightlifting/training in between same-day sessions (60% agreement).
- No additional care between sessions (60% agreement).
- Avoid big meals in between sessions (60% agreement).
- Don't do any vigorous back exercises two days before the test (55% agreement).
- Come back at the same time of the day (50% agreement).
- Don't undergo any physical demanding activity involving the back (45% agreement).
- Avoid swimming and scrubbing your back (40% agreement).
- Wear the same clothes for the next session (30% agreement).

14. If there are MULTIPLE MEASUREMENT SESSIONS in the same and/or different days, what procedures do you recommend to ensure the participant is oriented in the SAME POSITION as the previous measurement?

    (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)

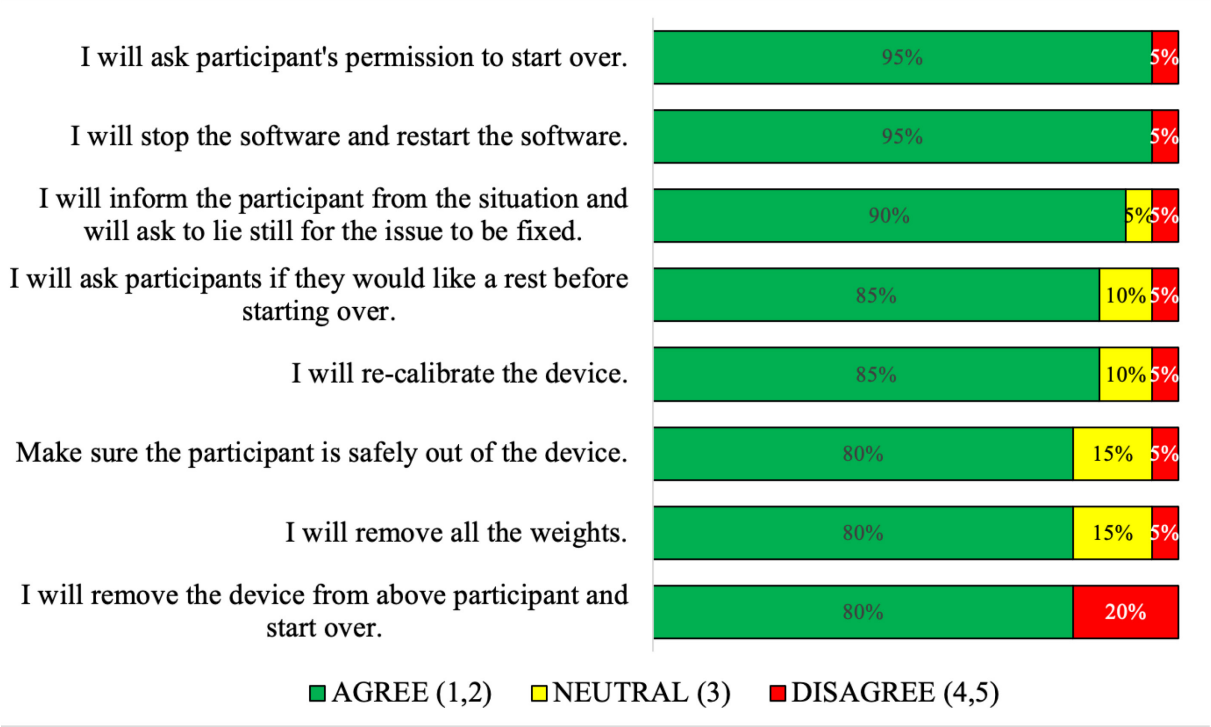


- Keep the reference points intact (95% agreement).
- Have a standardized examination table with markings that could be used to align participants in a reproducible manner (95% agreement).
- Take a photo with the consent of the participant (90% agreement).
- Use a permanent marker (particularly for S1) to ensure the starting position of measurement is the same (85% agreement).
- Put a band aid/ adhesive tape on top of the marked "x" spot so you don't lose it for the next visit (85% agreement).
- Measure the trajectory distance between T12 and S1 for the follow-up sessions (75% agreement).
- Take notes on the position of the patient (head, arms, legs) (65% agreement).
- Since the testing plinth has a hole, the participant will always align at approximately the same distance from the cephalic end of the plinth (65% agreement).

- Tape on the table and on the floor to ensure the same position of equipment and person on the table (65% agreement).
- Participants should feel just as comfortable as before (65% agreement).

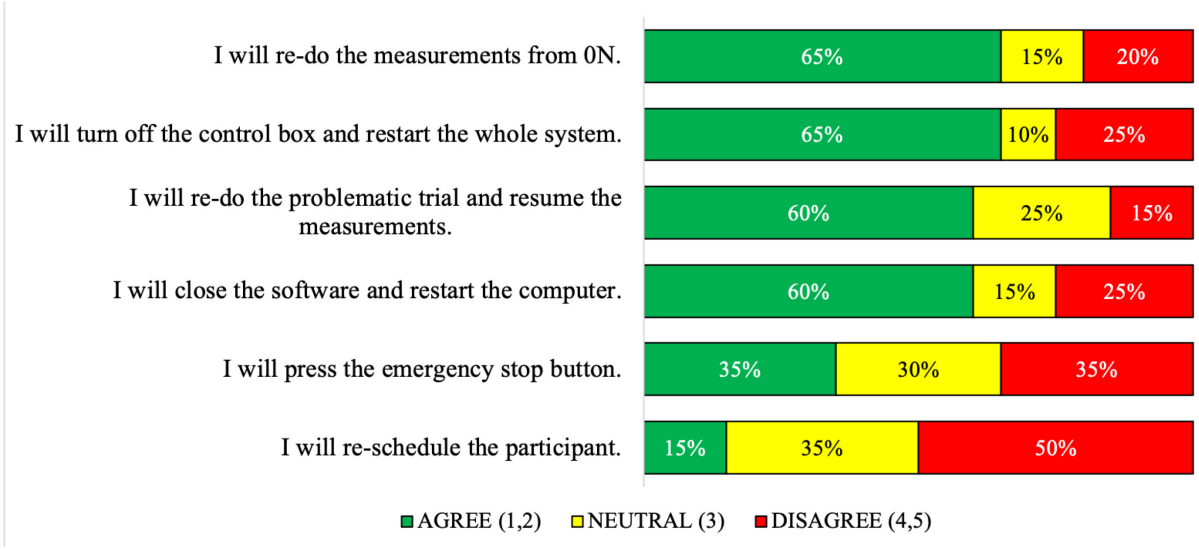
15. How would you handle a situation where the SOFTWARE PROGRAM CRASHES in the middle of a test for any reason?

    (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)



- I will ask participant's permission to start over (95% agreement).
- I will stop the software and restart the software (95% agreement).

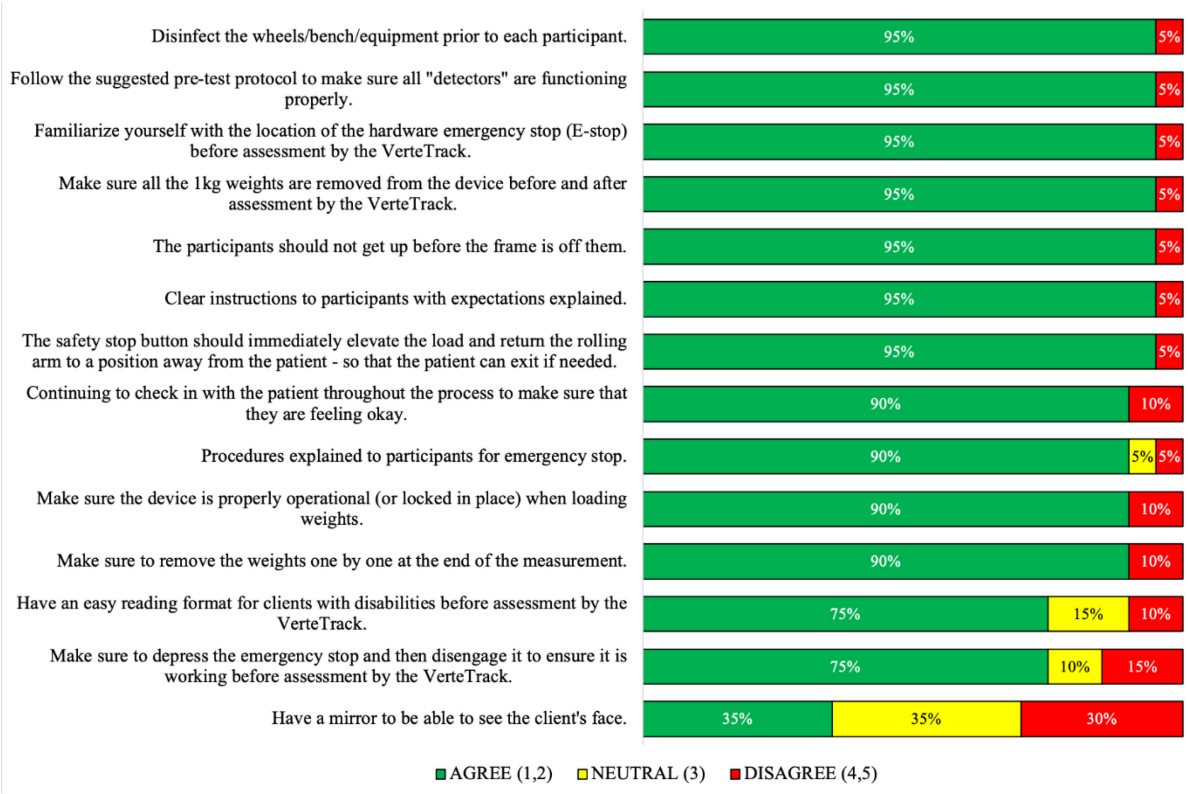
- I will inform the participant from the situation and will ask to lie still for the issue to be fixed (90% agreement).
- I will ask participants if they would like a rest before starting over (85% agreement).
- I will re-calibrate the device (85% agreement).
- Make sure the participant is safely out of the device (80% agreement).
- I will remove all the weights (80% agreement).
- I will remove the device from above participant and start over (80% agreement).



- I will re-do the measurements from 0N (65% agreement).
- I will turn off the control box and restart the whole system (65% agreement).
- I will re-do the problematic trial and resume the measurements (60% agreement).
- I will close the software and restart the computer (60% agreement).
- I will press the emergency stop button (35% agreement).
- I will re-schedule the participant (15% agreement).

- Software program crashes are less likely to be related to the control box issue. Therefore, turning off the computer or control box will be my last resort (new, no bar graph).
- My actions depend on the severity of the crash. For example, if I have to recalibrate the trajectory, I will have to recollect all trials (new, no bar graph).

16. Please indicate to what extent do you agree or disagree with the statements below regarding optimizing PARTICIPANT SAFETY when data are collected using the VerteTrack device. (1, strongly agree; 2, agree; 3, neither agree or disagree; 4, disagree; 5, strongly disagree)



- Disinfect the wheels/bench/equipment prior to each participant (95% agreement).
- Follow the suggested pre-test protocol to make sure all "detectors" are functioning properly (95% agreement).
- Familiarize yourself with the location of the hardware emergency stop (E-stop) before assessment by the VerteTrack (95% agreement).
- Make sure all the 1kg weights are removed from the device before and after assessment by the VerteTrack (95% agreement).
- The participants should not get up before the frame is off them (95% agreement).
- Clear instructions to participants with expectations explained (95% agreement).
- The safety stop button should immediately elevate the load and return the rolling arm to a position away from the patient - so that the patient can exit if needed (95% agreement).
- Continuing to check in with the patient throughout the process to make sure that they are feeling okay (90% agreement).
- Procedures explained to participants for emergency stop (90% agreement).
- Make sure the device is properly operational (or locked in place) when loading weights (90% agreement).
- Make sure to remove the weights one by one at the end of the measurement (90% agreement).
- Have an easy reading format for clients with disabilities before assessment by the VerteTrack (75% agreement).
- Make sure to depress the emergency stop and then disengage it to ensure it is working before assessment by the VerteTrack (75% agreement).
- Have a mirror to be able to see the client's face (35% agreement).

- I will raise the plinth when not testing to make sure it will not drop if it malfunctions (new, no bar graph).

Appendix I. VerteTrack Operations Manual

1. GENERAL INFORMATION

This chapter describes general information about the device including its intended test population, operational requirements, and different components.

1-1. Test population

The VerteTrack device is a biomedical device designed to measure posteroanterior mobility of the thoracic and lumbar spine. This device may only be operated by trained individuals who are familiar with human spine anatomy and within a research study having ethical approval.

1-2. Physical Description

The VerteTrack device is a software-driven mechanical device that incrementally applies force to acquire, process, and display stiffness data.

The device is composed of 6 major components:

1. A cube-shaped aluminum frame (Width 1080 mm × Height 1090 mm × Length 1510 mm) with lockable casters. **Figure 1-1 (A)**

2. A custom-made wheel apparatus consisting of a vertical rod, two rolling wheels, and a laser pointer. **Figure 1-1 (B)**

3. A stepping motor system along X, Y, and Z axes with built-in encoders (resolution = 0.007 mm) (National Instruments, USA). **Figure 1-1 (A)**

4. A string potentiometer (resolution = 0.020 mm, TE Connectivity, USA).

5. Custom-made LabVIEW software used to operate the device (National Instruments, USA). **Figure 1-1 (A)**

6. Two emergency stop buttons (software and hardware) allowing the operator to stop the measurements at any time.

Note: **Always** use the software emergency stop first if needed before using the hardware stop. This is because the software stop will withdraw the wheels from the participant which is the desired way to abort testing. If the hardware stop is used, power to the motors is killed which will result in the current weight of the wheels being suddenly dropped, possibly on to the participant.

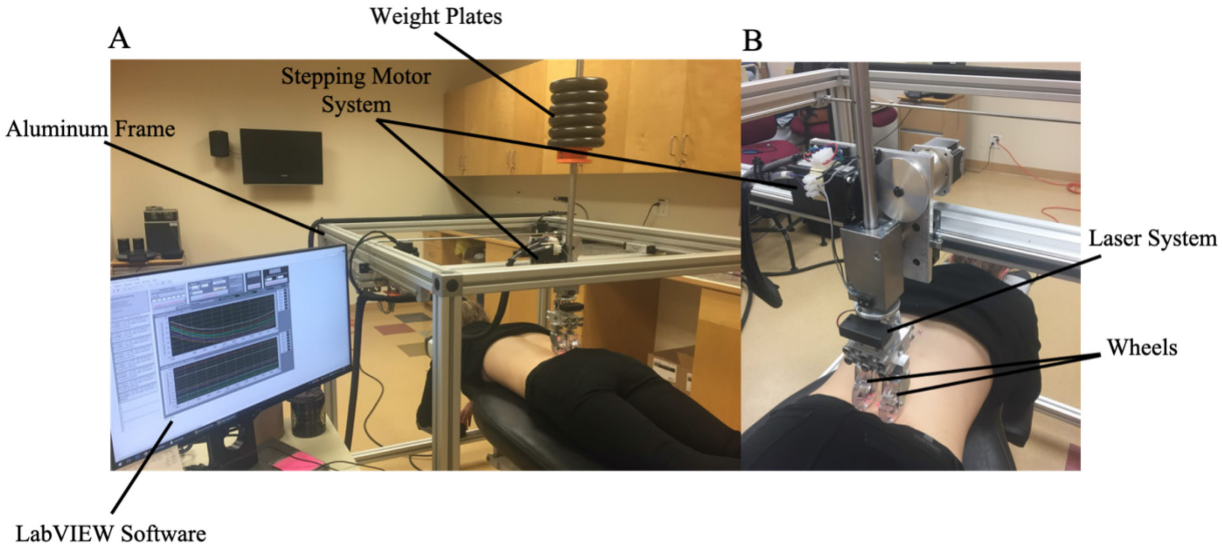


Figure 1-1. The VerteTrack Device: (A) The device components; (B) Roller Apparatus

1-3. Operational Summary

In brief, the movement of the load-bearing wheels is mapped out as a horizontal trajectory (XY) by manually moving the wheels to trace a path on the participant's spine with the laser system. Following that trajectory, a pre-selected vertical load is continuously applied by the wheel

apparatus while the resulting tissue deformation is measured by a wire string potentiometer. The string potentiometer measures the vertical position of the wheel apparatus and provides real-time feedback to the control system. The additional load can then be added to the wheels in increments of 10 N (Maximum = 60 N) using weight plates. Signals from the encoders that record the positions of the motors are collected by customized LabVIEW software at a collection rate of 200 Hz. The resulting force-deformation profile of the targeted segments (e.g., lumbar segments) can then be used to generate estimates of bulk stiffness along the trajectory.

1-4. Clinical Applications

- Thoracic spine
- Lumbar spine

1-5. Inclusion Criteria

- The participant's ability to tolerate a load of at least 40 N while in prone lying.
- Body mass index (BMI) under 40 for ease of locating spinous processes which define the wheel trajectory.

1-6. Exclusion Criteria

- Pregnancy.
- Skin lesions, infections, or open wounds in the vicinity of where the wheels will follow their trajectory (i.e., the skin surface of the lumbar and/or thoracic spine).
- An inability to lie in the prone position (e.g., severe deformities to spine or limbs, static tremor, uncontrolled epilepsy, etc.).

- Serious spinal pathology (e.g., spinal tumor, fracture, infectious disorder, osteoporosis, or other bone demineralizing condition, etc.).
- An inability to sustain held, active expiration (functional residual capacity) for at least 10 seconds.
- An inability to follow instructions (e.g., those with dementia or participants who may move during the test (e.g., age under 18)).
- A head, neck, or thoracoabdominal surgery within the last 6 months.
- Unstable spondylolisthesis.
- Unstable and/or acute disc herniation or injury.
- People who do not feel comfortable with the VerteTrack procedure.
- Unstable heart conditions.

2. Device Operation

This chapter will describe practical considerations for VerteTrack users. Operators will gain a fundamental appreciation of participants' needs, recording technique, device operation, feedback training, and data management. The purpose of this chapter is to ensure the correct and safe operation of the VerteTrack device (hardware, software, and mechanical).

2-1. Pre-testing Checklist

Before operating the device, make sure to:

- Remove each weight from the device.
- Familiarize yourself with the location of the hardware emergency stop (E-stop). Before using the device, test the emergency stop by depressing and then disengaging it. Ensure the E-stop is disengaged before use – this is a common reason why motors may not operate when the machine is first operated – someone has depressed the E-stop.
- Never step on or run over the device's cables which are custom-made and fragile.
- Wipe the wheels with a proper disinfectant (e.g., soapy water, alcohol).
- Ensure there are no cracks or other damage to the wheels before they are used. Inspect the wheels' surfaces for cracks and feel for cracks with fingertips as well.

Caution: **Do not** use wheels if they are cracked, damaged, or broken.

2-2. Device Preparation

- Lower the plinth completely using the plinth control. **Figure 2-1**

- Roll the frame away from the plinth. **Figure 2-2**
- Turn on the device computer and log in.
- Turn on the device control box. **Figure 2-3**
- Launch the device software located on the desktop (Swiper Control V07.1 Compiled APPLICATION.EXE.). **Figure 2-4**

Note: The device control box must always be turned on before launching the software. If the computer software was running before the control box was turned on, close the software, turn off the control box, then turn on the control box, and restart the software.

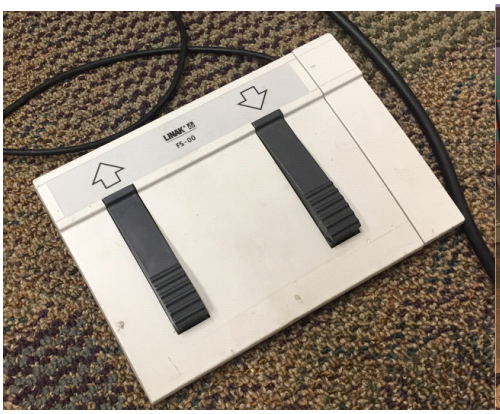


Figure 2-1. Plinth Control (example)

Figure 2-2. Frame away from the plinth.

Figure 2-3. Turn the control box on

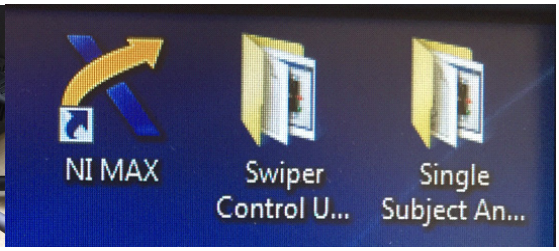


Figure 2-4. Launch Program

2-3. Check Limit Switches and Jogs

- Click the Check Limit Switches button in the software. **Figure 2-5**
- Depress each of the 6 limit switches and ensure the software light turns on for each sensor. Make sure the correct light turns on for the correct sensor.
- Click the Stop Checking button to end checking the limit switches. **Figure 2-6**
- At the top of the software, use the jog buttons to move the device in all 6 directions. **Figure 2-7**

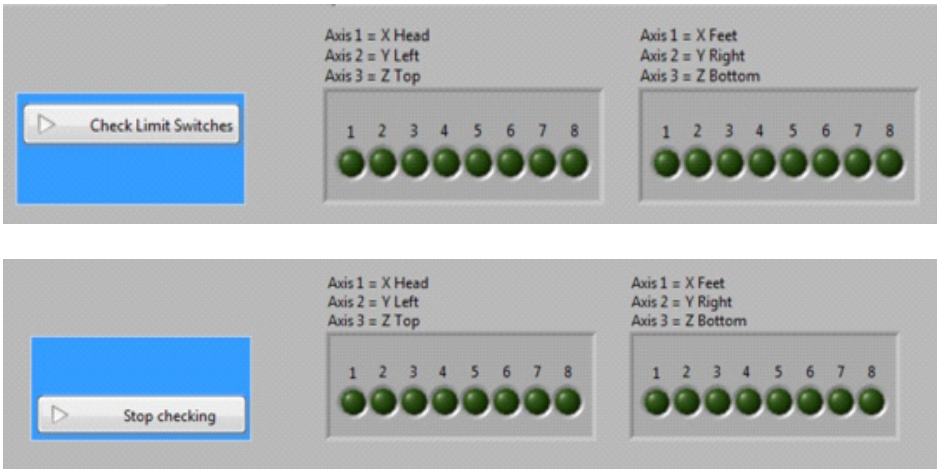


Figure 2-5. Limit switches in the software

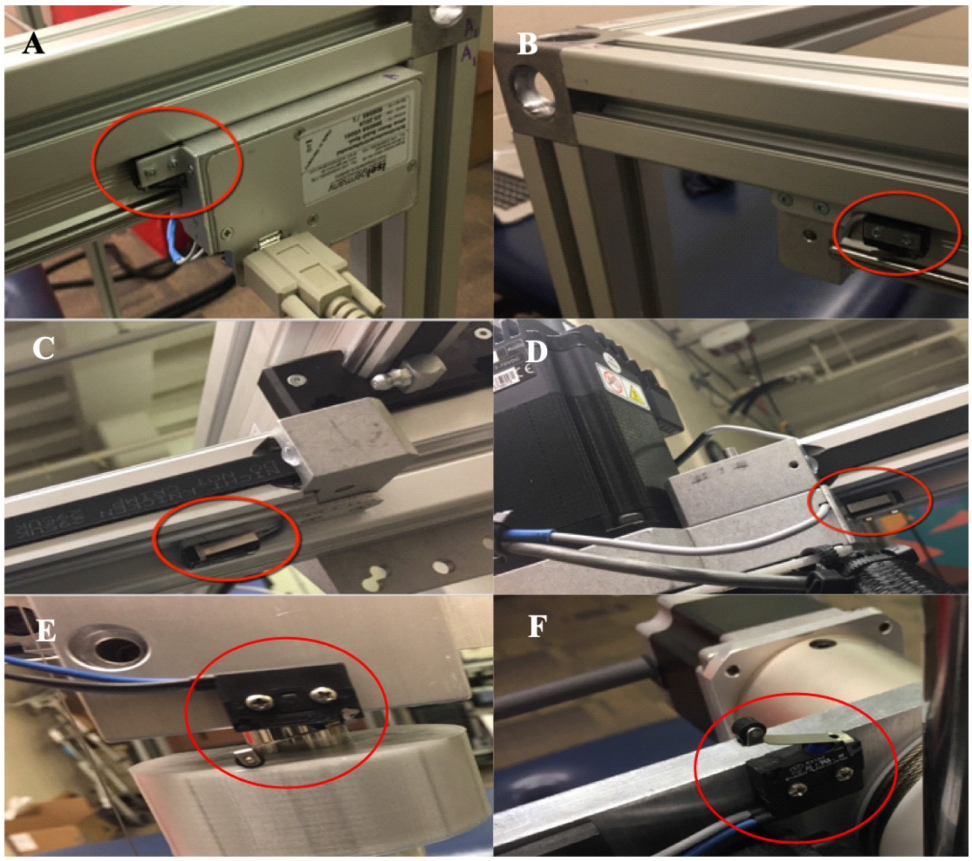


Figure 2-6. Mechanical limit switches: A) X-axis head, B) X-axis feet, C) Y-axis left, D. Y-axis right, E) Z-axis bottom, F) Z-axis top

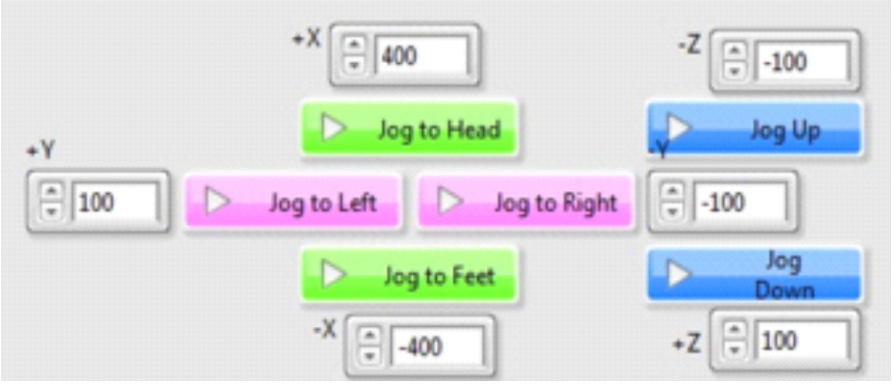


Figure 2-7. Jog buttons at the top of the software

2-4. Home the Device

- Before homing each axis, make sure that each direction of travel is clear of cables or other obstacles. Only run one homing direction at a time.
- Home the X-axis. **Figure 2-8**
    - Observe the process and watch for any cable impingement.
    - Wait for homing to complete – do not start any other processes
- Home the Y-axis. **Figure 2-8**
    - Observe the process and watch for any cable impingement.
    - Wait for homing to complete – do not start any other processes
- Home the Z-axis. **Figure 2-8**
    - Observe the process and watch for any cable impingement.
    - Wait for homing to complete – do not start any other processes

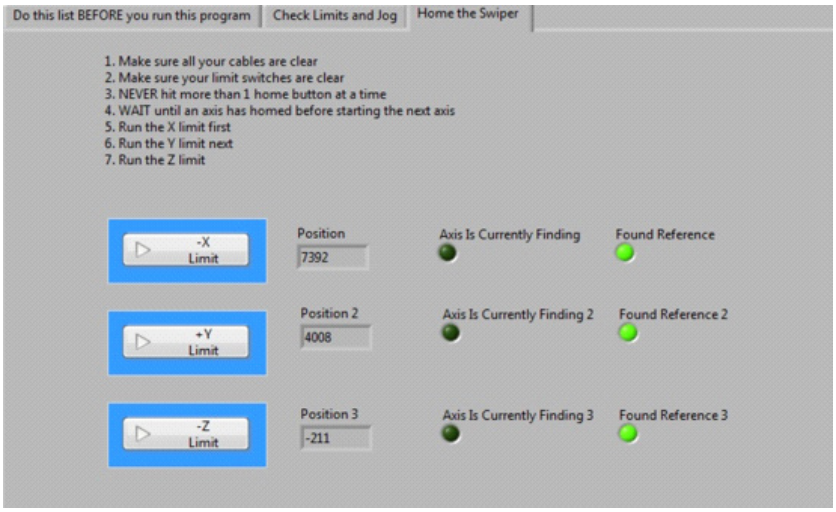


Figure 2-8. Home the device

2-5. Participant Welcoming Activities (Participant Testing, Page 1)

- Complete any forms for the study as much in advance as possible.
- Before testing begins, ensure the participant has gone to the restroom.
- Familiarize the participant with the procedure (See Chapter 3).

2-6 Determining Wheel Size and Inter-Wheel Distance

- For most participants, use the largest diameter wheels (3.0 inches).
- Before the participant is under the machine, examine their back and measure the width of the paraspinal tissues. **Figure 2-9**

Note: The wheels **should not** run on the spinous, nor too close to the spinous to avoid pulling the skin down over the spinous itself. Ideally, the wheels run over the top of the paraspinal tissues and above the lamina/facets. The wheels should be set no wider than the paravertebral tissues.

- Select the inter-wheel distance that best matches the distance below, or the next biggest size. For most adults, 1.125 inches (29mm) will suffice.
- If necessary, remove wheels with supplied wrench and change position.
- **If changing the wheels, or the inter-wheel distance, be careful not to twist wheel housing. This will potentially break the string potentiometer.**

Table 2-1. Wheel position with the corresponding distance between inner wheel face

Wheel position	**Distance between inner wheel face**	
Position 1, closest 2 holes	0.625 inches	16 mm
Position 2	1.125 inches	29 mm
Position 3	1.625 inches	41 mm
Position 4, furthest 2 holes	2.125 inches	54 mm

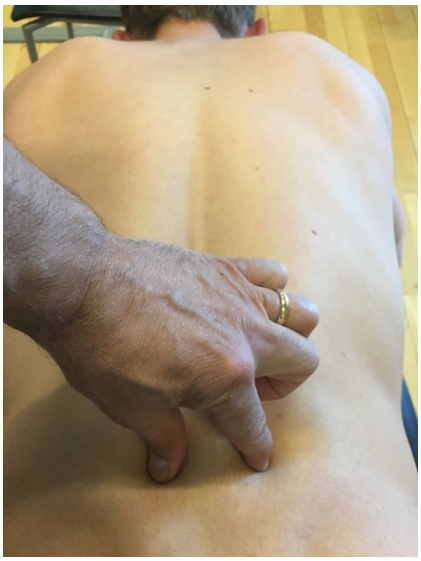


Figure 2-9. Gauging the width of spinous tissues

2-7. Mark the Participant (Participant Testing, Page 1)

- Place the participant in prone position on the plinth. Add foam roller under their legs.
- If using palpation, use a standardized palpation procedure based on anatomical landmarks (e.g., count up from the sacral base and down from T12/ribs).

- Place hands on iliac crests, identify the L4 spinous process, place a mark on the skin, go down towards the sacrum, identify the L5 spinous process, go up towards the thoracic vertebrae, identify each spinous process.
- Mark the starting position with an "x". **Figure 2-10**
- Confirm with diagnostic ultrasound if available.

- If possible, use ultrasound to begin with for better accuracy in identifying spinous processes.

Figure 2-10. Mark the participant

2-8. Positioning the Wheels over the Test Area (Participant Testing, Page 1)

- Clean the wheels for this participant.
- Make sure the plinth is at the lowest setting
- Roll the frame back over the participant so the middle of the frame is aligned T7.
- **Lock the wheels of the frame.**
- Tell the participant to **remain still** for the rest of the testing session.

- **Make sure to keep the arm position of participants the same – whatever you decide.** If possible, it is preferred to have the arms resting on supports that are part of the plinth itself.
- In software, click the Kill XY Axes Button.
    - If one of the axes becomes locked and cannot move, you likely ran into the limit switch. Kill axes again and move the wheels manually to the middle of the frame. **Figure 2-11**
- With the axes being killed, move the device manually so that the wheels are above the **highest point of the test** area (most posterior).
    - Lumbar spine – Likely the base of sacrum but check T/L region.
    - Thoracic spine – Apex of kyphosis.
    - **Ensure there is enough vertical travel space for the wheels to test the most posterior part of the participant's back.**
- **Only now, raise the plinth** until the highest point on the participant is 3 cm from the wheels. **Figure 2-1**

Note: **From this point on, do not change the plinth height for this participant.**

- Without changing the plinth height or moving the frame, move the wheels to the landing site by positioning the laser over the center of the "X".

- Make sure that the wheels are aligned in the direction of travel on the skin before running each trial.

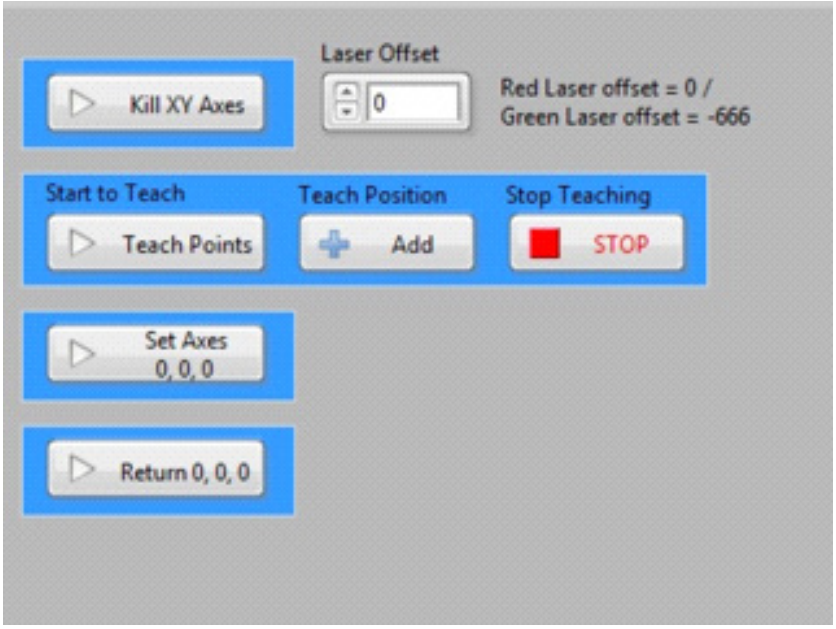


Figure 2-11. Kill Axes, Teach Points, Set Axes 000, Return 000

2-9. Teach FWD Trajectories (Participant Testing, Page 1)

- Turn on the laser.
- Without changing the plinth height or moving the frame, move the wheels to the Landing Site by positioning the laser over the center of the "X".

Note: When placing the laser, make sure to do so when the participant breathes out. The location of the skin markings can change quite a bit at different parts of the respiratory cycle.

- Click the Teach Points button. **Figure 2-11**
- Click the Add Point button immediately to capture 0, 0, 0 at the Landing Site. **Figure 2-11**
- Now move the laser manually to the next point and click the ADD A POINT button.

- Repeat this until all the points have been added including the “X” of the Lifting Point.
- When you are done, press the Stop Teaching button.

2-10. Select the Wheel Spacing and Wheel Size

- From the chart, pick the correct entry that describes the wheel spacing and wheel size. If using the largest wheel and second position, select the second choice (3.0 ICE, 1.125).
- Once you select the correction wheel diameter/position setting, press SET FILENAME.

2-11. Name FWD and REV Trajectory Files

- You will now be asked to name the FWD trajectory file. Use the participant's name or ID (whichever is appropriate) and include the letters FWD in the trajectory name. Save to the Desktop.
- You will now be asked to name the REV trajectory file.
    - Always save to the Desktop.
    - Use the participant’s name or participant ID (whichever is appropriate) and include the letters FWD in the trajectory name.
    - To avoid any startle, **let the participant know that as soon as you enter the rev filename, the device will move to the 0, 0, 0 position.**
- Click the Set Axes to 0,0,0 button. **Figure 2-11**
- Click the Return to 0,0,0 button. **Figure 2-11**
- Turn off the laser.

2-12. Load the Filenames (Participant Testing, Page 2)

- In the Data Collection Tab of the software, Load the filename of the FWD file in the appropriate text box. Do this by clicking on the file folder icon beside the text box. **Figure 2-12**
- Load the filename of the REV file in the adjacent text box. Do this by clicking on the file folder icon next to the text box.
- In the Participant File Name box, use the file folder icon to load a pre-existing trajectory file from the desktop. This will ensure that the correct path is loaded to the desktop. Then, change the filename in the textbox to follow the study's file naming protocol.

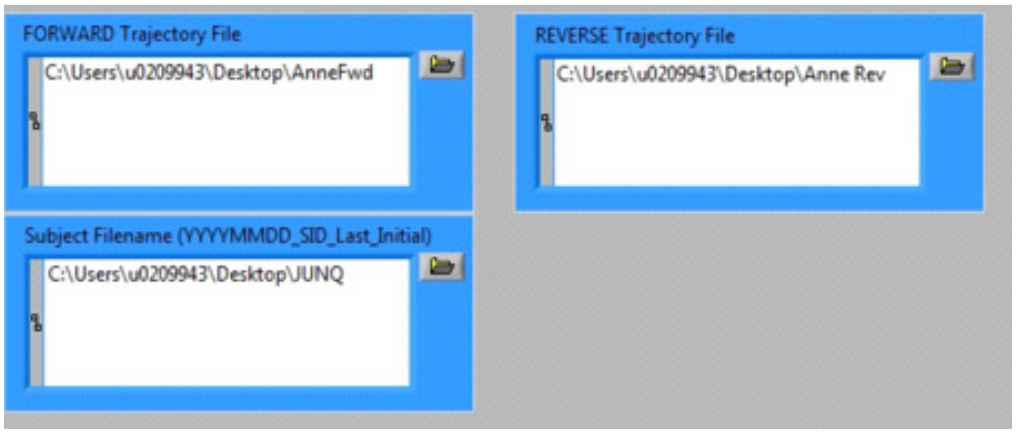


Figure 2-12. Load Filenames

2-13. Run the FWD Trial (Participant Testing, Page 2)

- Familiarize yourself with the Software Stop button and the Hardware E-stop. **Figure 2-13**
- Re-emphasize that the participant is to remain still during data collection.

- Add a 1 kg weight if needed (always add the additional 1 Kg weights to the device one at a time).
- Click the Start Trial button. **Figure 2-14**
- Review breathing with the participant if needed.
- Tell the participant that you are moving the wheels of the device down to touch their back.
- Using the newly appearing jog buttons (not the ones at the top of the screen), jog wheels down onto the participant and add enough cable slack (approx. 5 extra jogs down).
- Some participants with hyper-lordosis may require more than 5 extra jogs down.
- Align the wheels in the direction of travel before each trial if needed. **Figure 2-15**
- Look at the laser from the same angle to ensure it is lined up perfectly before each trial.
- When the wheels are fully touching the participant's skin and they are settled under the applied mass, check if the participant is ready, then start the breathing protocol and run the trajectory. **Figure 2-14**
- When the test is completed, the wheels will lift on their own.
- Now you can instruct the participant to breathe again.

Note:

- Do not leave the weight on the participant any longer than necessary. As soon as it settles on their skin and enough slack is provided, begin breathing and run the trajectory.
- Always consider the first round as a practice round to let the participant feel the testing procedure.

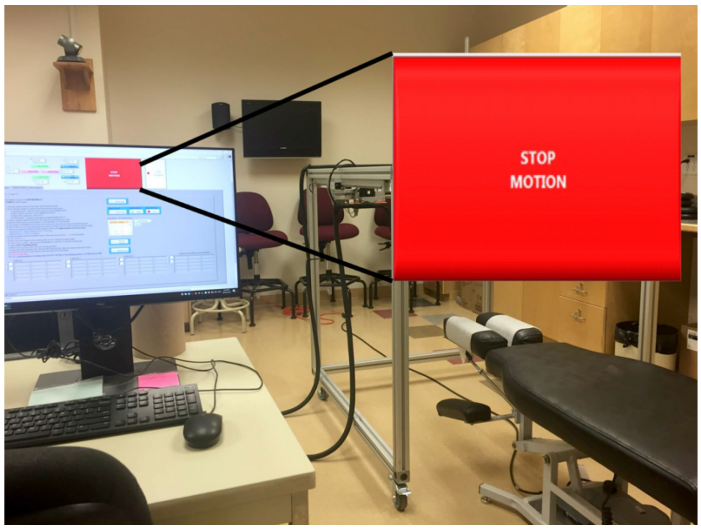


Figure 2-13. Software Emergency Button

Figure 2-14. Start trial, Trial Jog, Run Trajectory

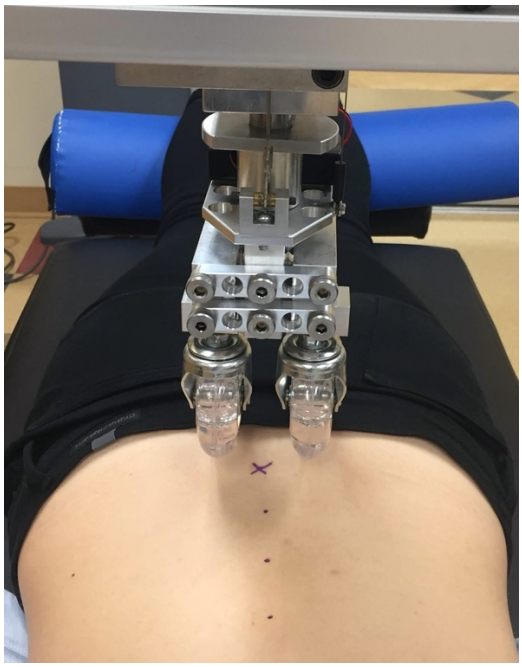


Figure 2-15. Make sure the wheels are aligned in the direction of travel before each trial

2-14. Return the Wheels by Running the REV Trial (Participant Testing, Page 2)

- Run the trajectory and the reverse trajectory will complete without the wheels touching the participant.
- Make sure the laser goes back to the reference point prior to subsequent runs.

2-15. Check the Trial, Save the Trial (Participant Testing, Page 2)

- Watch the participant during the trial to see if they breathe, move, perform a Val Salva or voluntarily contract their muscles. This may be a sign of discomfort.
- Check the data graph of the trial to ensure there were no breathing or movement artifacts in the collected data.

- Update the filename as appropriate. **Figure 2-16**

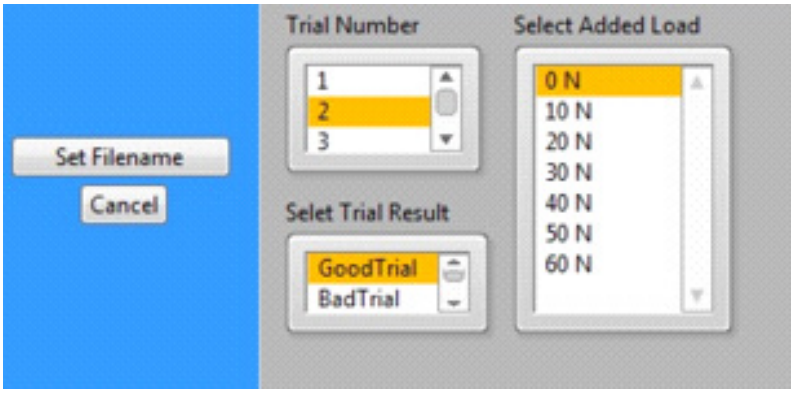


Figure 2-16. Revise trial filename

2-16. A Good/Bad Trial Definitions

- A good trial is a trial where the wheels follow the curvature of the spine without deviating sideways, and which does not cause discomfort to the participant.
- A good trial is one in which the participant is relaxed, does not move, and holds their breath for the entire trial.
- A good trial is one with consistent data collected for a single participant
- A bad trial is one with irregular change in the trajectory line.
- A bad trial is one where the participant breathes, moves, contracts muscles or contains any other event that would not occur in a normal test.
- If the wheels did not move smoothly and were not continuously pointed forward, it is a bad trial.

2-17. Instructions for Operator to Ensure a Good Trial

- Make sure you provide enough cable slack before the wheels move on the back.
- Make sure to align the wheels in the direction of travel.
- Communicate with the participant what you expect from them and give them regular feedback.
- Look for participant movement, breathing, and tonicity.
- Don't forget to check the graphical readout of the data after each trial.
- Double-check the data collected and make sure you have all your files before letting the participants leave. Repeat testing as needed.
- Look at the data graphs after each trial to make sure the data looks appropriate.
- Check the displacement values.
- If you noticed something different with the process, mark it as a bad trial.
- It is highly recommended that the plinth on which the participant is positioned has armrests to rest the arms in prone position.

2-18. Collecting the Next Trial (Participant Testing, Page 2)

- Make sure to align the wheels and the wheel housing before the next trial.
- Make sure the participant does not move between the trials (to talk to you, scratch, etc.).
- Go to Run the START TRIAL step.

2-19. End Data Collection with this Participant

- Tell the participant that the testing has ended but caution the participant not to get up until you tell them.
- Lower the participant plinth.
- Unlock frame wheels and roll frame away from the plinth.
- Help the participant to sit on the plinth remembering they may have difficulty after being prone. Ensure they do not bump their head on anything when arising from the plinth.
- Make sure the participant sits first on the plinth and is not dizzy before they stand up.
- Stand by to assist the participant as they stand.
- Thank the participant and give them any instructions for further testing etc.
- Now that the participant has left, remove weights from the device **one at a time**.

2-20. Testing Another Participant (After Participant)

- Go back to "Welcome Activities", Tab 1, Page 1

2-21. If You Are Finished with this Session (After Participant)

- Stop the program using the Big white software button. **Figure 2-17**
- **Always make sure there is no mass on the system before turning off** the computer or the control box.
- Turn off the device control box.
- Shut down the computer so those without passwords cannot operate the device.

- Note: If there is a long time between testing participants, your computer settings may put the system to sleep and the wheels may lose power and drop. If you are leaving the computer for extended hours, shut down the system completely.

Figure 2-17. Stop program button

2-22. If You Stopped the Software and Want to Start Again (After Participant)

- If you tried to use the software and it did not respond, make sure that you put any buttons pressed back to their original status.
- Then, press the white arrow key in the upper left corner of the program – it should turn black to tell you the program is running.

Note:

- Never leave mass on the device unattended or after the device is powered down.
- Always turn the control box off after testing. It is not meant to be left on.

2-23. Troubleshooting

If the software program crashes, you probably pushed the wrong button at the wrong time.

- Inform the participant about the situation and ask them to lie still for the issue to be fixed.
- Never turn off the computer or the control box when the participant is being tested.
- Ask participant's permission to start over.
- Ask participants if they would like a rest before starting over. If so, remove all the weights, then remove the device from above the participant and make sure the participant is safely out of the device.
- You may have to stop the program and restart:
    - Click the small Stop Sign icon in the upper left of the program screen. **Figure 2-18**
    - The program is now stopped.
    - If the button you pressed is still pressed (it is not the usual color), press it to return it to the normal state.
    - Start the program again by clicking the Arrow Icon to the left of the Stop Sign Icon in the upper left screen. **Figure 2-18**
    - Re-do the problematic trial and resume the measurements.

- Depending on where you were in the program when it crashed, you might have to remove all the weights and go to the Participant Testing, Page 1 and return the wheels to the 0, 0, 0 position by clicking Return to 0, 0, 0. If the wheels do not return to the starting point for that participant, you will have to click the Kill XY Axes button and reposition the wheels to the Landing Site using the laser. You can then either teach new trajectories or use existing trajectories:
    - If using existing trajectories, skip the Teach Points section and just click the Set to 0, 0, 0, and Return to 0, 0, 0 buttons. Note, whenever you use existing trajectories,

the participant must be exactly in the same position as when the trajectories were first collected.

  - If using new trajectories, you have to recollect all trials from 0N again.

- You may need to re-calibrate the device depending on the severity of the crash. In this case, you will have to recollect all trials.
- You may need to turn off the computer or control box as your last resort.

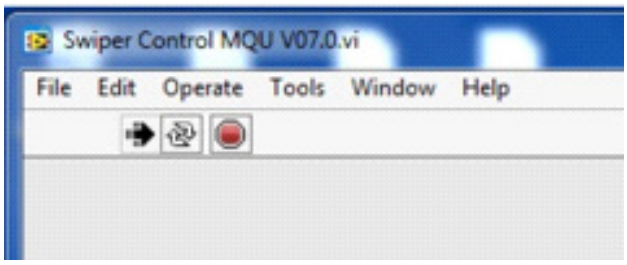


Figure 2-18. Program Start/Stop Buttons

3. Practical Considerations

Participants may be anxious about having weight placed on their back, as well as curious about the VerteTrack device and why spinal stiffness measurements are being taken. They also may not be familiar with anatomy and the rationale for recording stiffness data and so will not intuitively understand how their ability to follow instructions may affect data capture. Outcomes may be enhanced by performing some preliminary steps to familiarize the participant with clinical procedures. In this chapter, step-by-step instructions and examples are provided to improve the experience of both the operator and participant when using VerteTrack.

3-1. Participant Briefing (pre-test)

1. Briefly explain the rationale and goals for the spinal stiffness measurements and feedback training in terms appropriate to the participant's level of understanding.
2. Briefly explain the operation of the VerteTrack device
    a. Show participant the orientation video.
    b. Show the device to the participant in person, pointing out the different parts and explaining their function to help them further understand the process.
    c. Practice breathing protocols with the participant before beginning the measurements.
3. Explain to the participant what he or she can expect to feel and do during the session.

3-2. Example Remarks to Participant (pre-test)

“Low back pain is a major cause of pain, disability, and increased healthcare costs all over the world. Studies have shown that people with back pain seem to have more back stiffness. It has been also shown that with treatment, this stiffness may improve in some people but not others. We think that measuring spinal stiffness will help us to learn more about people with back pain and possibly teach us more about how we can help them. We can measure spine stiffness faster and more accurately using an instrument called a VerteTrack. This instrument includes a frame, a wheel system with a laser attached to it, and computer software.

We want you to be able to relax when we take the measurements so, please take your time to go to the restroom before we begin. During the test, you will lie face down and we may ask you to disrobe/change as necessary to expose your back (the test area) sufficiently. Gowns will be provided, but you will not need to change if you wear comfortable clothes that can be moved to expose your waistline. You will be also required to empty your front and back pockets of all objects and we ask that you remove your belt and glasses during the test. Once you are laying on the table, a researcher will touch your back in several places and then mark your skin with an ink pen. These marks are used to determine where we should measure the stiffness of your back. Next, to familiarize you with the instrument and the process we will perform a practice measurement.

During the test, you will feel the wheels roll over the skin of your back. This will feel a bit like a massage. There will be no needles, electric shocks, or other unexpected sensations. The only thing you should feel is the wheels touching your skin. We will add one light weight at a time to the machine which will simply make you feel like the wheels are adding more pressure to your back.

This should not hurt at any time and most people are comfortable with having up to 6 weights placed on the machine. However, if you feel too much pressure, let us know and we will stop the measurements. During the test, we will guide you to hold an exhaled breath for about 10 seconds while data is recorded. We will let you know when you can return again to normal breathing. When we record the data, we ask that you do not move or talk to us, unless you need or want, to stop the test. Between measurements, we may ask you a few questions to ensure you are feeling well. When speaking to us, we ask that you do not move your head or body. We will be monitoring the results closely and if necessary, we may need to repeat one or more measurements to ensure high-quality data capture.

Remember that you can ask to stop immediately at any time if you feel you cannot or do not want to continue with your participation in the trial or in taking the measurements. You will not be penalized for doing so. Do you have any questions or concerns? …Shall we proceed?"

3-3. Participant Briefing (during the test)

1. Make sure that the participant follows all the instructions (breathing, movement, relaxation).
2. Monitor any participant reactions to testing and be sure to ask how participants feel when adding the weight plates.

3-4. Example Remarks to Participant (during the test)

“Relax your back and abdominal (stomach) muscles and try to remain still for the duration of the test (e.g.,15 min). We also ask that you try to keep your arms in the same position for the duration of the test because moving your arms can cause changes in the tone of muscles in your back which can affect measurements. Breathe normally unless instructed to hold your breath or exhale.

The wheels are on your skin now and we’re going to start the test. Please take a deep breath in, now breathe out and hold it…hold…hold…keep holding…now you can breathe again. (Repeat with the participant through each addition of weight).

This is going to be the last measurement for now, but we ask that you not move or get off the table until we instruct you to. We will check the data and if everything looks good or if we need to repeat a measurement, we will let you know. Please wait for our instructions before you move away from the table”

3-5. Participant Briefing (post-test)

1. Make sure that the participant gets up safely.
2. Explain to the participant what he or she can expect to feel and do after the session.

3-6. Example Remarks to Participant (post-test)

After checking all the graphs: “We are done with our measurements. When it is time, I will help you sit up from the plinth.”

After removing the device assist the participant while saying: “Slowly get up from the table. If you need assistance, let us know. You may experience some dizziness; if so, please advise us of it and stay seated for a few minutes before standing up.”

Discharging the participant: “You may feel slight discomfort or back stiffness or show signs of mild skin irritation following our examination today. These things are normal and any discomfort you feel should improve on its own within two days, however, if it does not subside or you feel worried do not hesitate to contact the principal investigator.”

3-7. Participant Briefing (between tests)

1. Depending on what is being investigated, you may need to control for different activities including exercise, food intake, hydration levels (e.g., abdominal contents, gas, delayed onset muscle soreness, etc.).
2. Ask participants to refrain from treatments to the spine between sessions unless this treatment is a part of the experiment being conducted.
3. If the two testing sessions are a few hours apart, ask the participant to go for a walk and remind them to use the restroom before the second session starts.

3-8. Example Remarks to Participant (between tests)

“During the time between your test sessions, try to avoid any activities that could change the strength of your back muscles, for example, heavy weightlifting or personal training and performing sit-ups or other core strengthening exercises). You may continue with your usual activities (at home and at work) but we ask that you do not start any new exercises between your

visits to our lab. If you take medication like muscle relaxants or pain killers, try to take the medication after the assessment."

3-9. Recommendations for Operators to Reach the Same Position over Multiple Measurement Sessions

- Use a permanent marker (particularly for S1) to ensure the starting position for each measurement is the same.
- Ask the participant not to remove the ink marks on purpose or to add other marks to the test area.
- Use a standardized examination plinth with armrests and a face hole for prone positioning.
- Take a photo of the markings on the back (if allowed in your ethics procedure) to aid in repositioning the wheels on future testing dates.
- Consider using a band-aid/ adhesive tape to cover the marked "x" spot so you don't lose it for the next visit.
- Measure the trajectory distance.

3-10 Summary of Recommendations for Operators to Optimize Participant Safety

- Familiarize yourself with the location of the hardware emergency stop (E-stop) before performing patient assessment using the VerteTrack.
- The safety stop button should immediately elevate the load and return the rolling arm to a position away from the participant to allow them to exit the device if needed.

- Before participant testing **only**, check the E-stop by depressing the emergency stop and then disengage it to ensure it is working.
- Ensure participants know that they can request the testing b stopped at any time and without penalty.
- Use an easy reading instruction format for clients with disabilities before assessment by VerteTrack.
- Disinfect the wheels/bench/equipment prior to seeing each participant.
- Provide clear instructions to participants with expectations explained.
- Follow the suggested pre-test protocol to make sure all "detectors" are functioning properly.
- Make sure the device is properly operational (or locked in place) when loading weights.
- Continue to check in with the participant throughout the process to make sure that they are feeling okay.
- Make sure all of the weights are removed from the device before and after assessment by the VerteTrack.
- Make sure to remove the weights one by one at the end of the measurement.
- Do not allow the participant to get up before the frame has been moved away from them.

Appendix J. Ethics Approval Form_ Delphi study

Notification of Approval

Date:	August 17, 2020
Study ID:	Pro00102734
Principal Investigator:	Maliheh Hadizadeh Bajestani
Study Supervisor:	Gregory Kawchuk
Study Title:	**Development of a standard protocol for spinal stiffness measurement using a novel loaded rolling wheel system: a Delphi study**
Approval Expiry Date:	Monday, August 16, 2021

Approved Consent Form:	Approval Date	Approved Document
	8/17/2020	Informed Consent

Thank you for submitting the above study to the Research Ethics Board 2. Your application, including the following, has been reviewed and approved on behalf of the committee;

- Recruitment Emails (8/11/2020)
- Questionnaire 1 (8/11/2020)

Any proposed changes to the study must be submitted to the REB for approval prior to implementation. A renewal report must be submitted next year prior to the expiry of this approval if your study still requires ethics approval. If you do not renew on or before the renewal expiry date, you will have to re-submit an ethics application.

Approval by the Research Ethics Board does not encompass authorization to access the staff, students, facilities or resources of local institutions for the purposes of the research.

Approval by the Research Ethics Board does not encompass authorization to recruit and/or interact with human participants at this time. Researchers still require operational approval as applicable (eg AHS, Covenant Health, ECSD etc) and where in-person interactions are proposed, institutional and operational requirements as outlined in the Resumption of Human Participant Research - June 24, 2020 must be met.

Sincerely,

Dr. Ubaka Ogbogu, LLB, BL, LLM, SJD
Chair, Research Ethics Board 2

Note: This correspondence includes an electronic signature (validation and approval via an online system).

Appendix K. Screening Form

Participant ID (PID) _____

Initial Screening Date _____

Gender ☐ Male ☐ Female

Inclusion Criteria

Current age (Eligible age range is 18-60 yrs. old) _____

Currently feel limited by at least 20% of daily function because of low back pain (Try to determine if enough disability for possible ODQ score 20% or higher: If YES-potentially eligible; If NO-NOT eligible)

☐ Yes ☐ No

Exclusion Criteria

Currently pregnant (If YES - NOT eligible) ☐ Yes ☐ No

Prior surgery to the lumbosacral spine (If YES - NOT eligible) ☐ Yes ☐ No

Currently receiving mind-body or exercise treatment for LBP from a healthcare provider (e.g., No chiropractic, physical therapy, massage therapy, etc.) (If YES - NOT eligible)

☐ Yes ☐ No

Initial Screening Status

☐ Eligible – interested ☐ Eligible – not interested ☐ Ineligible

Screening Notes_____

Appendix L. Eligibility Form

Eligibility Date_____

Date Informed Consent Form (ICF) Signed_____

(If ICF not signed, leave field blank and add comments in Eligibility Notes field at the end of the form).

Inclusion Criteria

Current age (Eligible age range is 18-60 yrs. old) _____

Oswestry disability score > 20% (If NO - NOT eligible) ☐ Yes ☐ No

Pain between the 12th rib and buttocks with or without symptoms into one or both legs, which, in the opinion of the examiner, originate from the lumbar region. (If NO - NOT eligible)

☐ Yes ☐ No

Exclusion Criteria

Currently pregnant (If YES - NOT eligible) ☐ Yes ☐ No

Prior surgery to the lumbosacral spine (If YES - NOT eligible) ☐ Yes ☐ No

Currently receiving mind-body or exercise treatment for LBP from a healthcare provider (e.g., chiropractic, physical therapy, massage therapy, etc.) (If YES - NOT eligible)

☐ Yes ☐ No

Neurogenic sign of positive ipsi- or contra-lateral straight leg raise test < 45 degrees in either leg AND symptoms reproduced (If YES - Not eligible) ☐ Yes ☐ No

Right SLR (In degrees) _____

Left SLR (In degrees) _____

Other Neurogenic signs including any of the following: reflex, sensory, or strength deficit in a pattern consistent with lumbar nerve root compression (If YES - NOT eligible)

☐Yes ☐No

Any "red flags" of a potentially serious condition including:

- Numbness, tingling, or weakness (If YES - NOT eligible)
- Fever
- Night pain
- Pain that does not change with movement
- Weight loss
- Difficulty passing urine or having bowel movement
- Recent trauma or fracture
- History of cancer, prolonged corticosteroid use, systemic disease
- Drug abuse, immunosuppression, HIV

Eligibility notes: _____

Appendix M. Consent and Information Sheet_ University of Alberta

Optimization of Spinal Manipulative Therapy Protocols

Local Principal Investigator: Gregory N, Kawchuk, BSc, DC, MSc, PhD

Overall Principal Investigator: Julie M. Fritz, Ph.D., PT, ATC

Why am I being asked to take part in this research study?

We are inviting you to take part in this research study because you have low back pain. You are not eligible to take part in this research study if you have had back surgery in the past, if you know you are pregnant, or if you are currently receiving treatment for your back pain from a physical therapist, chiropractor, massage therapist or other type of medical provider. Please note that have consulted any of these health care practitioners allows you to participate but you cannot be undergoing care during the trial. Also, a researcher will evaluate your back and if you have any signs of a problem with the nerves in your back (i.e., "sciatica") or any indication that your back pain may be due to a problem such as an infection, fracture or cancer, you will not be eligible to participate.

Before you make a decision one of the researchers will go over this form with you. You can always ask questions if you need a better explanation or more information. You will be given a copy of this form to keep.

What is the reason for doing the study?

We are conducting this study to better understand how spinal manipulation affects people with back pain and how we can improve this treatment. The purpose of this experiment is to determine how spinal manipulation can reduce stiffness in the back and help back muscles function better. We are also examining what exercises can be added to spinal manipulation to make it more effective. This study is being sponsored by the National Institutes for Health.

What will I be asked to do?

If you decide to participate in this study, you will be asked to attend three sessions in Phase I (Figure 1 below). These three sessions will occur 2-3 days apart. During the first two Phase I sessions you will receive an evaluation and spinal manipulation treatment. The third session will include an evaluation only. Each of these sessions will last about 30 minutes.

At the third session of Phase I you will be randomly assigned to additional treatment to be provided over the next 3 weeks in Phase II. Randomization means that you will be put into a group by random chance. This means a computer will decide what additional treatment you get in Phase II, not the study investigators. There are eight possible treatment groups that you might be randomized to in Phase II. Phase II Treatment groups may include receiving additional spinal manipulation treatment and/or additional exercises designed to either strengthen back muscles or decrease spine stiffness. One group in Phase II will receive no additional treatment. Because there are 8 groups for Phase II, your chances of being in any one group is 1 in 8 or 12.5%. If you are in a group in Phase II that receives additional treatment, you will receive two treatment

sessions per week over the next 3 weeks for a total of 6 treatment sessions. Each session will last about 30 minutes. The total treatment period for this study, including both Phase I and Phase II, will last 4 weeks. The diagram on the next page shows the flow of the study from Phase I to Phase II with the different treatment options for Phase II. More details of the treatment options are also described below.

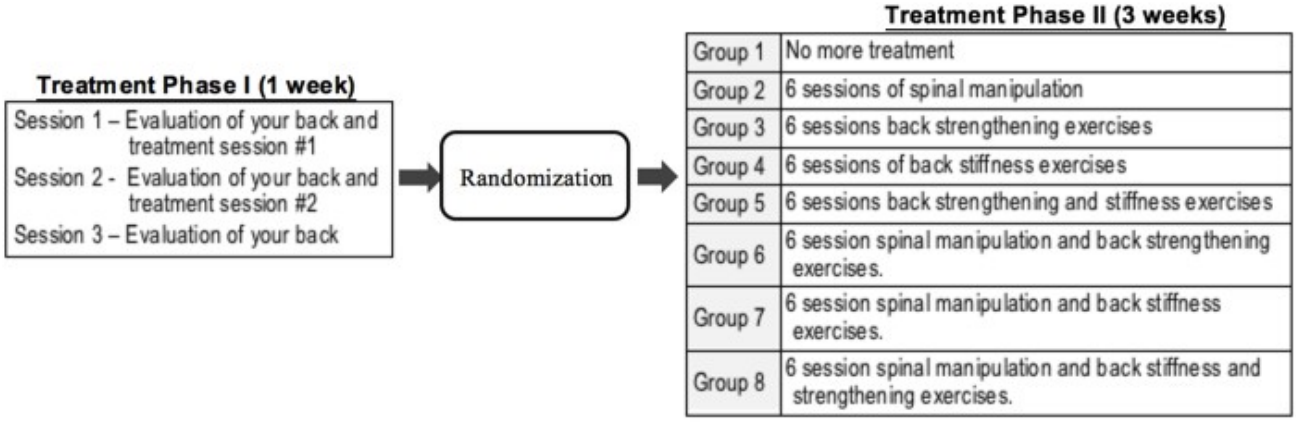


Figure 1.

Study examination procedures

If you agree to participate in this study, you will be provided with a secure web address to complete a form about your back pain. Completing this form means that you allow us to use its information. If you complete this form, but do not participate in the study, this information will not be kept. If more than 7 days pass between collecting this information and your first visit to the university, or something else happens to change your back pain (like a new injury), we may have to collect this information again.

We will then schedule at time for you to come to the university where you will be asked to sign a consent form to participate in the full study. If you consent to participate, we will ask you to complete several other surveys about your general medical history, how back pain affects

your activities, and also receive a physical examination to assess your back muscle function with ultrasound and spinal stiffness. In total, all the surveys will take approximately 10 minutes to complete, and the physical examination will take approximately an additional 20 minutes to complete. All of these procedures are used to check that you are eligible to participate in the study, and determine your level of pain and function before starting any treatment. These examination procedures will be repeated at the end of Phase I and the end of Phase II.

Ultrasound measurement procedures

As a part of each examination, we will use an ultrasound machine to measure the function of your deep stomach and back muscles. Ultrasound is a machine that transmits sound waves through the body and records the echoes as the sound waves move through different structures in the body. The echoes are transformed into images that can be viewed on a television screen. During the ultrasound measurements you will be asked to lie on your back or your stomach. A liquid gel will be placed on your skin to help transmit the sound waves. The ultrasound device will then be placed on your skin and you will be asked to perform three simple tasks while ultrasound measurements are taken. While lying on your back you will be asked to lift one leg off the table, and you will be asked to tighten your stomach muscles. While lying on your stomach you will be asked to lift one arm off the table. In each case, you may be asked to lift a small weight (~ 1-3kg) with your arm or leg.

Spinal stiffness measurement procedures

As a part of each examination, we will use a spinal stiffness machine to measure the stiffness in your back. The machine uses a tool that looks like a wheel to roll over your spine. During the

stiffness measurements you will be asked to lie on your stomach. The wheel will be lowered onto your lower back and will feel like someone pushing a tennis ball slowly on to the surface of your back. You will then be asked to hold your breath and you will feel this same pushing sensation as the wheel rolls up and down your back. We will roll the wheel up and down your back 10 times with each time adding with a little more weight to increase the pushing sensation into your back. At the end of the test, you will feel a push that is about the same as having a bag of sugar resting on your back. You will always have one minute to rest between stiffness tests. During the test, and afterward, we want you to tell us if there is any pain or discomfort During each stiffness test, you will be given a trigger to squeeze. If you want the stiffness test to stop, squeeze the trigger and the stiffness test will stop.

Spinal manipulation

During the first and second sessions of Phase I, and possibly during Phase II you will be given a spinal manipulation treatment ("cracking your back"). The manipulation will be performed by having you lay on your back. A researcher will gently bend and rotate your body. You may feel and/or hear a "pop" during the manipulation procedure. The researchers have had extensive training in this form of treatment and use it routinely in their practice. You may also discontinue the treatment if you feel discomfort at any time.

Spinal stiffness exercises

These exercises are designed to improve the flexibility of your spine and the muscles around your back. The exercises involve lying on your stomach or your back or being on all fours and moving in different directions to stretch the muscles around your back. You will be given a copy

of the exercises with pictures and instructions for when the exercises are to be completed at home. There may be 2-4 exercises done up to 10 times daily.

Back muscle strengthening exercises

These exercises are designed to improve the strength of the muscles around your back. The exercises involve lying on your back or side or being on all fours and contracting different muscles around your back. You will be given a copy of the exercises with pictures and instructions for the number of times the exercises are to be completed at home. There may be 2 exercises that are done up to 20 times daily.

You can ask to stop at any time if you feel you cannot continue or do not want to continue. If at any time you want to stop the testing or finish the study, let the investigator know immediately.

What are the risks and discomforts?

The risks associated with participation in this study are minimal. If you are uncomfortable with any questions asked or they cause you distress, you do not have to answer. You may experience muscle soreness in your trunk, arms, or legs from the examination. Based on our experience this type of soreness is common meaning that it occurs in 1% to 25% of participants. There is also a chance that the study treatment may be ineffective or could exacerbate your back pain. We have attempted to minimize this risk by having licensed healthcare professionals perform all study procedures and by ensuring that all participating healthcare providers have been thoroughly trained in the procedures to be used in this study. There are no known risks from the ultrasound measurements and it has been found safe to use over the abdominal region of pregnant women. It is also possible you may experience soreness after spinal manipulation or stiffness testing.

Based on our experience, this type of soreness is common meaning that it occurs in 1% to 25% of participants but also only lasts 24 hours. We have attempted to minimize this risk by allowing you to stop the testing if this soreness occurs.

You may also experience some side effects that cannot be predicted. Because of this, it is important that you tell one of the researchers listed below right away if you feel any unusual feeling or symptoms or if have any concerns. There may be risks in this study that are currently not known. If we find out anything new during the course of this research which may change your willingness to be in the study, we will tell you about these findings.

Number of participants

We expect to enroll 140 participants at the University of Alberta. We expect to enroll 280 participants total for this study at all sites.

What happens if I am injured because of this research?

If you become ill or injured as a result of being in this study, you will receive necessary medical treatment, at no additional cost to you. By signing this consent form you are not releasing the investigator(s), institution(s) and/or sponsor(s) from their legal and professional responsibilities."

What are the benefits to me?

You may experience health benefits from the treatment provided in the study but this cannot be guaranteed. We hope that the information we get from doing this study will help us better understand how to better measure spine so that we can use this measure in future tests.

Do I have to take part in the study?

Being in this study is your choice. If you decide to be in the study, you can change your mind and stop being in the study at any time without penalty. In the event of you opt out of the study once we have started collecting data, we will continue to use the data we have collected unless you don't want this to happen. You do not have to answer any questions that you are not comfortable with.

Will I be paid to be in the research?

You will receive a small token payment to help any costs associated with your involvement (e.g. parking). Specifically, you will be paid $25.00 for each assessment that you complete in this study. This includes a maximum of 4 assessments. Total compensation you will receive for participating is therefore $100.00 if each session is completed.

Will my information be kept private?

Your personal health record related to this study will be kept confidential. Any research data collected about you during this study will not identify you by name, only by a coded number. Your name will not be disclosed outside the research center. Any report published as a result of this study will not identify you by name. Any data shared between the study sites will be done so by encrypted technologies.

The health information collected in this study will be kept confidential unless release is required by law. All information will be used only for the research study. The researchers and

the Health Research Ethics Board may access your study records to monitor the research and verify the accuracy of study information.

In Canada, study information is required to be kept for 7 years. Even if you withdraw from the study, the information and data that is obtained from you for study purposes will not be destroyed. You have the right to check your health records and request changes if your personal information is incorrect.

A description of this clinical trial will be available on http://www.ClinicalTrials.gov, as required by law. This website will not include information that can identify you. At most, the website will include a summary of the results. You can search this website at any time.

Funding and conflict of interest

The study is being sponsored by the National Institutes of Health. The University of Alberta and the investigators conducting this study are getting money from the study sponsor to cover the costs of doing this study. You are entitled to request any details concerning this compensation from Dr. Greg Kawchuk. No University of Alberta researchers will receive direct compensation for conducting this study. There is a small chance that the results from the study could help commercialize processes used in this study to better identify and treat low back pain.

What if I have questions?

If you have concerns about your rights as a study participant, you may contact the Research Ethics Office at (780) 492-2615. This office has no affiliation with the study researchers.

Gregory Kawchuk	Professor in Physical Therapy	University of Alberta	780-492-6891
Julie Fritz	Professor in Physical Therapy	University of Utah	801-581-6297

CONSENT FORM

Title of Project: Optimization of Spinal Manipulative Therapy Protocols

Local Principal Investigator: Dr. Gregory Kawchuk

Overall Principal Investigator: Dr. Julie Fritz

_____

Part 2 (to be completed by the research subject):

	Yes	No
Do you understand that you have been asked to be in a research study?	☐	☐
Have you read and received a copy of the attached Information Sheet?	☐	☐
Do you understand the benefits and risks involved in taking part in this research study?	☐	☐
Have you had an opportunity to ask questions and discuss this study?	☐	☐
Do you understand that you are free to withdraw from the study at any time, without having to give a reason and without affecting your future medical care?	☐	☐
Has the issue of confidentiality been explained to you?	☐	☐
Do you understand who will have access to the information you provide?	☐	☐
Do you want the investigator(s) to inform your family doctor that you are participating in this research study? If so, give his/her name _____	☐	☐

Who explained this study to you?_____

I agree to take part in this study: YES ☐ NO ☐

Signature of Research Subject: _____

(Printed Name): _____ Date:

_____

I believe that the person signing this form understands what is involved in the study and voluntarily agrees to participate.

Signature of Investigator or Designee: _____ Date

_____

THE INFORMATION SHEET MUST BE ATTACHED TO THIS CONSENT FORM AND A COPY GIVEN TO THE RESEARCH SUBJECT

Appendix N. Consent and Authorization Document_ University of Utah

BACKGROUND

You are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with friends and relatives if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you volunteer to take part in this research study. We are conducting this study to better understand how spinal manipulation effects people with back pain and how we can improve this treatment. The purpose of this experiment is to determine how spinal manipulation can reduce stiffness in the back and help back muscles function better. We are also examining what exercises can be added to spinal manipulation to make it more effective. This study is being sponsored by the National Institutes for Health.

We are inviting you to take part in this research study because you have low back pain. You are not eligible to take part in this research study if you have had back surgery in the past, if you know you are pregnant, or if you are currently receiving treatment for your back pain from a physical therapist, chiropractor, massage therapist or other type of medical provider. Also, a researcher will evaluate your back and if you have any signs of a problem with the nerves in your back (i.e., "sciatica") or any indication that your back pain may be due to a problem such as an infection, fracture, or cancer, you will not be eligible to participate.

STUDY PROCEDURES

If you decide to participate in this study, you will be asked to attend three sessions in Phase I. These three sessions will occur 2-3 days apart. During the first two Phase I sessions you will receive an evaluation and spinal manipulation treatment. The third session will include an evaluation only. Each of these sessions will last about 30 minutes.

At the third session of Phase I you will be randomly assign to additional treatment to be provided over the next 3 weeks in Phase II. Randomization means that patients are put into groups by random chance. This means a computer will decide what additional treatment you get in Phase II, not the study investigators. There are eight possible treatment groups that you might be randomized to in Phase II. Phase II Treatment groups may include receiving additional spinal manipulation treatment and/or additional exercises designed to either strengthen back muscles or decrease spine stiffness. One group in Phase II will receive no additional treatment. Because there are 8 groups for Phase II, your chances of being in any one group is 1 in 8 or 12.5%. If you are in a group in Phase II that receives additional treatment, you will receive two treatment sessions per week over the next 3 weeks for a total of 6 treatment sessions. Each session will last about 30 minutes. The total treatment period for this study, including both Phase I and Phase II, will last 4 weeks. We will evaluate you one more time 3 months after you begin the study. The diagram on the next page shows the flow of the study from Phase I to Phase II with the different treatment options for Phase II. More details of the treatment options are also described below.

<table>
<tr><th>Treatment Phase I – 1 Week</th><td rowspan="9">1 Week

Post-Phase I Assessment
↓
Randomize to Phase II Group</td><th colspan="2">Treatment Phase II - 3 Weeks</th><td rowspan="9">4 Weeks

Post-Phase II Assessment</td><td rowspan="9">3 Months

Final Assessment</td></tr>
<tr><td rowspan="8">Enrollment
&
2 Treatment Sessions of Spinal Manipulation</td><td>Group 1:</td><td>No more treatment</td></tr>
<tr><td>Group 2:</td><td>6 sessions of spinal manipulation</td></tr>
<tr><td>Group 3:</td><td>6 sessions back strengthening exercises</td></tr>
<tr><td>Group 4:</td><td>6 sessions of back stiffness exercises</td></tr>
<tr><td>Group 5:</td><td>6 sessions back strengthening and stiffness exercises</td></tr>
<tr><td>Group 6:</td><td>6 session spinal manipulation and back strengthening exercises.</td></tr>
<tr><td>Group 7:</td><td>6 session spinal manipulation and back stiffness exercises.</td></tr>
<tr><td>Group 8:</td><td>6 session spinal manipulation and back stiffness and strengthening exercises.</td></tr>
</table>

Study Examination Procedures

If you agree to participate in this study, at your first session you will receive an examination that consists of you completing several questionnaires about your general medical history and how back pain affects your activities, and then you will receive a physical examination to assess your back muscle function with ultrasound and spinal stiffness. The questionnaires will take approximately 10 minutes to complete, and the physical examination will take approximately an additional 20 minutes to complete. The examination procedures will check to be certain that you are eligible to participate in the study and determine your level of pain and function before starting any treatment. These examination procedures will be repeated at the end of Phase I (after 1 week), at the end of Phase II (after 4 weeks) and after 3 months.

Ultrasound Measurement Procedures

As a part of each examination, we will use an ultrasound machine to measure the function of your deep stomach and back muscles. Ultrasound is a machine that transmits sound waves through the body and records the echoes as the sound waves move through different structures in the body. The echoes are transformed into images that can be viewed on a television screen. During the ultrasound measurements you will be asked to lie on your back or your stomach. A liquid gel will be placed on your skin to help transmit the sound waves. The ultrasound device will then be placed on your skin, and you will be asked to perform three simple tasks while ultrasound measurements are taken. While lying on your back you will be asked to lift one leg off the table, and you will be asked to tighten your stomach muscles. While lying on your stomach you will be asked to lift one arm off the table.

Spinal Stiffness Measurement Procedures

As a part of each examination, we will use a spinal indenter machine to measure the stiffness in your back. The indenter machine uses a tool that looks like a baton to press on your spine and record how stiff your spine is. During the stiffness measurements you will be asked to lie on your stomach. The tool will be pressed onto an area of your spine. This will feel like someone pushing their thumb slowly into your spine. At the end of the test, you will feel a push that is about the same as having a bag of sugar resting on your back. Three tests of stiffness will be done during each evaluation, for a total of 15 stiffness tests. During each stiffness test, you will be given a trigger to squeeze. If you want the stiffness test to stop, squeeze the trigger and the operator will stop the test. After the tool has been taken off your back, you rest for a few seconds.

Study Treatment Procedures

Spinal Manipulation

During the first and second sessions of Phase I, and possibly during Phase II you will be given a spinal manipulation treatment ("cracking your back"). The manipulation will be performed by having you lay on your back. A researcher will gently bend and rotate your body. You may feel and/or hear a "pop" during the manipulation procedure. The researchers have had extensive training in this form of treatment and use it routinely in their practice. You may also discontinue the treatment if you feel discomfort at any time.

Spinal Stiffness Exercises

These exercises are designed to improve the flexibility of your spine and the muscles around your back. The exercises involve lying on your stomach or your back or being on all fours and moving in different directions to stretch the muscles around your back. You will be given a copy of the exercises with pictures and instructions for the number of times the exercises are to be completed at home.

Back Muscle Strengthening Exercises

These exercises are designed to improve the strength of the muscles around your back. The exercises involve lying on your back or side or being on all fours and contracting different muscles around your back. You will be given a copy of the exercises with pictures and instructions for the number of times the exercises are to be completed at home.

RISKS

The risks associated with participation in this study are minimal. You may experience muscle soreness in your trunk, arms, or legs from the examination. Based on our experience this type of soreness is common meaning that it occurs in 1% to 25% of participants. There is also a chance that the study treatment may be ineffective or could exacerbate your back pain. We have attempted to minimize this risk by having licensed healthcare professionals perform all study procedures and by ensuring that all participating healthcare providers have been thoroughly trained in the procedures to be used in this study. There are no known risks from the ultrasound measurements, and it has been found safe to use over the abdominal region of pregnant women. There is a risk of soreness from the indenter machine. Based on our experience this type of soreness is common meaning that it occurs in 1% to 25% of participants. We have attempted to minimize this risk by allowing you to stop the testing if this soreness occurs.

UNFORESEEABLE RISKS

In addition to the risks listed above, you may experience a previously unknown risk or side effect.

BENEFITS

We cannot promise any benefits from your being in the study. You may benefit from participation by receiving spinal manipulation and exercises for your back. We hope that the treatments will help your back pain. However, this cannot be guaranteed. The information we get from this study may help us to treat future patients with low back pain better.

ALTERNATIVE PROCEDURES

If you decide not to take part in this study, there are a few alternative treatments or procedures available. You could be referred to physical therapy or a chiropractor and receive manual therapy, heat and/or cold therapy, ultrasound (deep heat), electrical stimulation, or different exercises for your low back pain. Alternatively, you could pursue specialty physician care, which may include the use of injections, acupuncture, surgical consultation, and/or other treatments for your back.

PERSON TO CONTACT

If you have questions, complaints or concerns about this study, you can contact Dr. Julie Fritz at (801) 587-2237. If you feel you have been injured as a result of participation, please call Dr. Julie Fritz at (801) 587-2237**.**

Institutional Review Board: Contact the Institutional Review Board (IRB) if you have questions regarding your rights as a research participant. Also, contact the IRB if you have questions, complaints or concerns which you do not feel you can discuss with the investigator. The University of Utah IRB may be reached by phone at (801) 581-3655 or by e-mail at irb@hsc.utah.edu.

Research Participant Advocate: You may also contact the Research Participant Advocate (RPA) by phone at (801) 581-3803 or by email at participant.advocate@hsc.utah.edu.

RESEARCH-RELATED INJURY

If you are injured from being in this study, medical care is available to you at the University of Utah Hospital, as it is to all sick or injured people. The University of Utah has not set aside any

money to pay the costs for such care. The University will work with you to address costs from injuries. Costs would be charged to you or your insurance company (if you have insurance), to the study sponsor or other third party (if applicable), to the extent those parties are responsible for paying for medical care you receive. Since this is a research study, some health insurance plans may not pay for the costs. By signing this consent form you are not giving up your right to pursue legal action against any parties involved with this research.

The University of Utah is a part of the government. If you are injured in this study, and want to sue the University or the doctors, nurses, students, or other people who work for the University, special laws may apply. The Governmental Immunity Act of Utah is a law that controls when a person needs to bring a claim against the government and limits the amount of money a person may recover. See sections 63G -7-101 to -904 of the Utah Code.

VOLUNTARY PARTICIPATION

Participation in this study is voluntary. If you decide not to participate there are no penalties and your health care will not be affected. If you decide to participate, you may discontinue participation at any time without any penalty. If you want to stop being in this study, please let the researchers know.

COSTS AND COMPENSATION TO PARTICIPANTS

There will be no cost to you for participation in this study. You will not be charged, nor will your insurance company be charged, for any test or visit that is completed for this study. You will be compensated for participation in this research. You will be paid $25.00 for each assessment that

you complete in this study. This includes a maximum of 4 assessments. Total compensation you will receive for participating is therefore $100.00 if each session is completed.

NUMBER OF PARTICIPANTS

We expect to enroll 140 participants at the University of Utah. We expect to enroll 280 participants total for this study at all sites.

AUTHORIZATION FOR USE OF YOUR PROTECTED HEALTH INFORMATION

Signing this document means you allow us, the researchers in this study, and others working with us to use some information about your health for this research study.

This is the information we will use and include in our research records:

- Demographic and identifying information like name, address and email address
- Related medical information about you like your medical history, current and past medications or therapies for your back, and information from physical examinations such as spine stiffness and muscle function measures.
- All tests and procedures that will be done in the study

How we will protect and share your information:

- We will do everything we can to keep your information private, but we cannot guarantee this. Study information will be kept in a secured manner and electronic records will be password protected. Study information may be stored with other information in your medical record. Other doctors, nurses, and third parties (like insurance companies) may be able to see

this information as part of the regular treatment, payment, and health care operations of the hospital. We may also need to disclose information if required by law.

- A description of this clinical trial will be available on http://www.ClinicalTrials.gov, as required by U.S. Law. This website will not include information that can identify you. At most, the website will include a summary of the results. You can search this website at any time.

- In order to conduct this study and make sure it is conducted as described in this form, the research records may be used and reviewed by others who are working with us on this research:
    - Members of the research team
    - The University of Utah Institutional Review Board (IRB), which reviews research involving people to make sure the study protects your rights.
    - Other academic research centers we are working with: University of Alberta
    - The study sponsor: The National Institutes of Health
- If we share your identifying information with groups outside of the University of Utah Health Sciences Center, they may not be required to follow the same federal privacy laws that we follow. They may also share your information again with others not described in this form. If you do not want us to use information about your health, you should not be part of this research. If you choose not to participate, you can still receive health care services at University of Utah Health Sciences Center.

What if I decide to Not Participate after I sign the Consent and Authorization Form?

You can tell us anytime that you do not want to be in this study and do not want us to use your health information. You can also tell us in writing. If you change your mind, we will not be able to collect new information about you, and you will be withdrawn from the research study. However, we can continue to use information we have already started to use in our research, as needed to maintain the integrity of the research.

This authorization does not have an expiration date.

CONSENT

I confirm that I have read this consent and authorization document and have had the opportunity to ask questions. I will be given a signed copy of the consent and authorization form to keep.

_____

Participant's Name

_____ _____

Participant's Signature Date

_____

Name of Person Obtaining Authorization and Consent

_____ _____

Signature of Person Obtaining Authorization and Consent Date

Appendix O. Ethics Approval Form_ University of Alberta

Health Research Ethics Board

308 Campus Tower
University of Alberta, Edmonton, AB T6G 1K8
p. 780.492.9724 (Biomedical Panel)
p. 780.492.0302 (Health Panel)
p. 780.492.0459

Date:	August 29, 2016
Study ID:	Pro00067152
Principal Investigator:	Gregory Kawchuk
Study Title:	Optimization of Spinal Manipulative Therapy Protocols
Approval Expiry Date:	Monday, August 28, 2017
Approved Consent Form:	Approval Date 8/29/2016 Approved Document UofA Information and Consent V3 CLEAN
Sponsor/Funding Agency:	University of Utah (This is a subgrant from a NIH Award)

RSO-Managed Funding:

Project ID	Project Title	Speed Code	Other Information
View RES0030216	Optimization of Spinal Manipulative Therapy Protocols		

Thank you for submitting the above study to the Health Research Ethics Board - Health Panel. Your application, including the following, has been reviewed and approved on behalf of the committee.

- Recruitment Poster V2 (8/24/2016)
- Study
  Forms V2

(8/26/2016)
NIH
Protocol
(7/28/2016)

A renewal report must be submitted next year prior to the expiry of this approval if your study still requires ethics approval. If you do not renew on or before the renewal expiry date, you will have to re-submit an ethics application.

Approval by the Health Research Ethics Board does not encompass authorization to access the patients, staff or resources of Alberta Health Services or other local health care institutions for the purposes of the research. Enquiries regarding Alberta Health Services approvals should be directed to (780) 407-6041. Enquiries regarding Covenant Health should be directed to (780) 735-2274.

Sincerely,

Anthony S. Joyce, Ph.D.
Chair, Health Research Ethics Board - Health Panel

Note: This correspondence includes an electronic signature (validation and approval via an online system).

Appendix P. Ethics Approval Form _ University of Utah

INSTITUTIONAL REVIEW BOARD
THE UNIVERSITY OF UTAH
75 South 2000 East Salt Lake City, UT 84112 | 801.581.3655 | IRB@utah.edu

IRB: IRB_00092127

PI: Julie Fritz

Title: Optimization of Spinal Manipulative Therapy Protocols

Date: 7/28/2016

Effective 7/28/2016, the above-referenced protocol is approved to begin the research procedures outlined in the University of Utah IRB-approved application and documents.

APPROVAL DOCUMENTATION

Review Type: Convened Board Review
Risk Level: Greater Than Minimal
Approval Date: 7/20/2016
Expiration Date: 7/19/2017 11:59 PM

DETERMINATIONS

- **Waiver/Alteration Determination:** The IRB has determined that the request for the **waiver of authorization** as described in this application is approved for this research under 45 CFR 164.512(i).

APPROVED DOCUMENTS

Informed Consent Document
Consent Draft July 26 Clean Copy

Surveys, etc.
Study questionnaires and surveys

Grant Application
Specific aims and research strategy from R33

Recruitment Materials, Advertisements, etc.
recruitment flyer

Other Documents
Santillo CITI document
Santillo CITI document

ONGOING SUBMISSIONS FOR APPROVED PROJECTS

- **Continuing Review:** The research protocol must be re-reviewed and re-approved prior to the expiration date via the continuing review application: http://irb.utah.edu/submitapplication/reviews/index.php
- **Amendment Applications:** All changes to the research application, protocol, or approved documents must be submitted and approved prior to initiation: http://irb.utah.edu/submitapplication/amendments.php
- **Report Forms:** The research must adhere to the University of Utah IRB reporting requirements for unanticipated problems and deviations: http://irb.utah.edu/submit-application/forms/index.php
- **Final Project Reports for Study Closure:** The research application must be closed with the IRB once the research activities are complete: http://irb.utah.edu/submit-application/final-projectreports.php

Click IRB_00092127 to view the application and access the approved documents.

Please take a moment to complete our customer service survey. We appreciate your opinions and feedback.

Appendix Q. Oswestry Disability Index (ODI) Form

Please see the instructions below for completing the questionnaire.

This questionnaire has been designed to provide information about how your back pain has affected your ability to manage in everyday life. Please answer every question by selecting the one answer that best describes your condition today. We realize you may feel that two statements may describe your condition, but please select the one which most closely describes your current condition.

The participant must have a valid Informed Consent date in the Eligibility Form before using this form.

Pain Intensity

☐ I can tolerate the pain I have without pain medication.

☐ The pain is bad, but I can manage without having to take pain medication.

☐ Pain medication provides me complete relief from pain.

☐ Pain medication provides me with moderate relief from pain.

☐ Pain medication provides me with little relief from pain.

☐ Pain medication has no effect on my pain.

Personal Care (washing, dressing, etc.)

☐ I can take care of myself normally without causing increased pain.

☐ I can take care of myself normally, but it increases my pain.

☐ It is painful to take care of myself and I am slow and careful.

☐ I need help but I am able to manage most of my personal care.

☐ I need help every day in most aspects of my care.

☐ I do not get dressed, wash with difficulty and stay in bed.

Lifting

☐ I can lift heavy weights without increased pain.

☐ I can lift heavy weights, but it causes increased pain.

☐ Pain prevents me from lifting heavy weights off the floor, but I can manage if the weights are conveniently positioned (ex. on a table).

☐ Pain prevents me from lifting heavy weights, but I can manage light to medium weights if they are conveniently positioned.

☐ I can lift only very light weights.

☐ I cannot lift or carry anything at all.

Walking

☐ Pain does not prevent me from walking any distance.

☐ Pain prevents me from walking more than 1 mile.

☐ Pain prevents me from walking more than 1/2 mile.

☐ Pain prevents me from walking more than 1/4 mile.

☐ I can only walk with crutches or a cane.

☐ I am in bed most of the time and have to crawl to the toilet.

Sitting

☐ I can sit in any chair as long as I like.

☐ I can only sit in my favorite chair as long as I like.

☐ Pain prevents me from sitting for more than 1 hour.

☐ Pain prevents me from sitting for more than ½ hour.

☐ Pain prevents me from sitting for more than 10 minutes.

☐ Pain prevents me from sitting at all.

Standing

☐ I can stand as long as I want without increased pain.

☐ I can stand as long as I want but increases my pain.

☐ Pain prevents me from standing more than 1 hour.

☐ Pain prevents me from standing more than 1/2 hour.

☐ Pain prevents me from standing more than 10 minutes.

☐ Pain prevents me from standing at all.

Sleeping

☐ Pain does not prevent me from sleeping well.

☐ I can sleep well only by using pain medication.

☐ Even when I take pain medication, I sleep less than 6 hours.

☐ Even when I take pain medication, I sleep less than 4 hours.

☐ Evens when I take pain medication, I sleep less than 2 hours.

☐ Pain prevents me from sleeping at all.

Social Life

☐ My social life is normal and does not increase my pain.

☐ My social life is normal, but it increases my level of pain.

☐ Pain prevents me from participating in more energetic activities (ex. sports, dancing etc.).

☐ Pain prevents me from going out very often.

☐ Pain has restricted my social life to my home.

☐ I have hardly any social life because of my pain.

Traveling

☐ I can travel anywhere without increased pain.

☐ I can travel anywhere but it increases my pain.

☐ My pain restricts travel over 2 hours.

☐ My pain restricts my travel over 1 hour.

☐ My pain restricts my travel to short necessary journeys under 30 minutes.

☐ My pain prevents all travel except for visits to the doctor/therapist or hospital.

Employment / Homemaking

☐ My normal homemaking/job activities do not cause pain.

☐ My normal homemaking/job activities increase my pain, but I can still perform all that is required of me.

☐ I can perform most of my homemaking/job duties, but pain prevents me from performing more physically stressful activities (ex. lifting, vacuuming).

☐ Pain prevents me from doing anything but light duties.

☐ Pan prevents me from doing even light duties.

☐ Pain prevents me from performing any job or homemaking chores.

Raw Score _____

Questions answered_____

Adjusted score (Percentage) _____

Appendix R. The Fear-Avoidance Beliefs Questionnaire (FABQ)

Here are some of the things other patients have told us about their pain. For each statement please check the number from 0 to 6 to indicate how much physical activities or work affects or would affect your back pain.							
	0 Completely Disagree	1	2	3 Unsure	4	5	6 Completely Agree
Physical activity makes my pain worse.							
Physical activity might harm my back.							
I should not do physical activities which (might) make my pain worse.							
I cannot do physical activities which (might) make my pain worse.							
My pain was caused by my work or by an accident at work.							
My work aggravated my pain.							
My work is too heavy for me.							
My work makes or would make my pain worse.							
My work might harm my back.							
I should not do my regular work with my present pain.							
I do not think that I will be back to my normal work within 3 months.							

Appendix S. UWCAP & UWPRSE

Please rate how confident you are that you can do the following things at present, despite the pain. To indicate your answer, please select one answer per row. How confident are you that...					
	Not at all				Very much
You can get necessary work done in spite of your TYPICAL pain (if you don't work outside of home consider household work or unpaid work)?					
You can keep your pain from interfering with your social life?					
You can manage your pain during your daily activities?					
You can accomplish most of your goals in life in spite of your pain?					
You can cope with your pain in most situations?					
You can maintain an active lifestyle in spite of your pain?					
You can do the things you most want to do in spite of your pain?					
You can socialize with friends in spite of your pain?					
You can do most of your household chores in spite of your pain?					
You can have a fulfilling life in spite of your pain?					
In the past 7 days, how often did you have the following thought when you were in pain? Please respond to each item selecting one answer per row.					
	Never				Always
I can't stand my pain anymore.					
Because of my pain, my life is never going to get any better.					
I will lose everything because of my pain.					
I will never be able to do many of the things I enjoy because of my pain.					
Because of my pain, I will be in a bad mood for the rest of my life.					
My pain overwhelms me.					
Because of my pain, my life is terrible.					
Because of my pain, something really bad is going to happen to me.					
In the past 7 days, how often...					
Could you only focus on how bad your pain feels?					
Did your pain completely fill up your mind?					