

2013-06-14

# Derivation of a Clinical Decision Guide in the Diagnosis of Cervical Facet Joint Pain

Schneider, Geoff Martin

---

Schneider, G. M. (2013). Derivation of a Clinical Decision Guide in the Diagnosis of Cervical Facet Joint Pain (Doctoral thesis, University of Calgary, Calgary, Canada). Retrieved from <https://prism.ucalgary.ca>. doi:10.11575/PRISM/25559

<http://hdl.handle.net/11023/753>

*Downloaded from PRISM Repository, University of Calgary*

UNIVERSITY OF CALGARY

Derivation of a Clinical Decision Guide in the Diagnosis of Cervical Facet Joint Pain

by

Geoff Martin Schneider

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES  
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE  
DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MEDICAL SCIENCES

CALGARY, ALBERTA

MAY, 2013

© Geoff Martin Schneider 2013

## **Abstract**

Neck pain is a common problem presenting to health care professionals in a primary care setting. In particular, the facet joint has been implicated as a primary source of pain in 36% - 67% of people with persistent neck pain. The most internationally accepted criterion standard for the diagnosis of facet joint pain is comparative controlled injections of local anaesthetic onto the sensory nerve of the target facet joint. This procedure is invasive, costly, and is often associated with lengthy wait-times. Clinicians often apply findings from the patient history and physical examination prior to referral of their patients for diagnostic facet joint blocks. The diagnostic utility of the clinical examination has raised controversy in the literature.

The purpose of this thesis was to derive a clinical decision guide in the diagnosis of cervical facet joint pain. Our research revealed that the manual spinal examination, palpation for segmental tenderness, and extension-rotation tests possess moderate to excellent intra-rater and inter-rater reliability. Our research indicated that positive findings on all three clinical tests might be helpful for clinicians when attempting to identify patients who may respond positively to diagnostic facet joint blocks. Of interest, the manual spinal examination and palpation for segmental tenderness test may be useful screening procedures, as they possessed high sensitivity and low negative likelihood ratios. Future research, in independent samples, is needed to validate our findings or generate other decision guides that are optimal for clinicians to deliver high quality care for their patients with persistent neck pain.

## Preface

The chapters of this thesis include two papers that have been published in peer-reviewed journals. The coauthors of these chapters have all granted permissions for these chapters to be included in this thesis. Permission to reprint has also been obtained for each of these manuscripts. Chapter 2 was published in *Manual Therapy* under the following citation: Schneider G, Jull G, Thomas K, Salo P. Screening of patients suitable for diagnostic cervical facet joint blocks – a role for physiotherapists. *Manual Therapy Journal* 2012; 17(2): 180-183. My contribution to this article was the review of the literature, concept generation, and writing of the manuscript.

Chapter 3 was published in the *Archives of Physical Medicine and Rehabilitation* under the following citation: Schneider G, Jull G, Thomas K, Smith A, Emery C, Faris, P, Schneider K, Salo P. Intra-rater and inter-rater reliability of select clinical tests in patients referred for diagnostic facet joint blocks in the cervical spine. *Arch Phys Med Rehab* 2013. In press. DOI: 10.1016/j.apmr.2013.02.015. My role in this paper was to formulate the research question, design the study, collect the data, analyse the data and write the manuscript. Supplementary documents for Chapter 3 can be found in Appendices A, B, and C.

## Acknowledgements

The work contained in this thesis would not have been possible had it not been for the support and mentorship provided by my supervisors, Dr. Paul Salo and Dr. Ken Thomas. Their expertise, knowledge and support have made this process an exceptional learning experience and time for personal and professional growth. They are leaders in their field and have afforded me many opportunities to learn from their example. Thank you for your dedication to this process.

This work would also not have been possible without the support of my supervisory committee, Dr. Carolyn Emery, Dr. Bevan Frizzell and Dr. Peter Faris. Their willingness to share their expertise in methodology, statistics, and interventional pain medicine was highly appreciated and facilitated the development of this research. I would also like to thank Kiran Pander for her amazing support and organization during the process of this degree. Thank you to Heather Scherer, JoAnna Herring, Yvette Andreas, and Scott Sherman for their research support and clinical assistance. Without their help, this work would not have been completed in a timely manner!

A special thanks is necessary to the Canadian Institutes of Health Research for their generous support via their scholarship funding. I also wish to thank the Calgary Orthopaedic Research and Education Fund, the Alberta Spine Society, and LifeMark Health for their funding support of this research. A special thank you is warranted for Prof. Gwendolen Jull, whose perseverance, expertise, and mentorship in this area of research is admirable and motivational. Thank you to the radiologists,

radiology technicians, and support staff at EFW Radiology. Their dedication to this research was outstanding and without them, this work would not be possible.

I have had the opportunity to work with a variety of professors, fellow students, and clinicians throughout the course of my doctoral work. I learned from each and every one of them and value the time that we have spent critically appraising research and brainstorming for future studies. In particular, I want to thank Ashley Smith for his commitment to evidence-based practice and his desire to better our profession. He sets an example of what patient care should resemble, and deserves acknowledgment of this.

I would also like to thank my wife Kathryn and our children Emma, Myla, and Kaden for their love, support, and encouragement during this busy time. Thank you for allowing me the time to delve into this project and for supporting me throughout the journey. Thank you to Kathryn for being my 'rock' and shining example of how you can be a wonderful spouse, parent, and professional all in one. Our discussions regarding this clinical research have had a significant impact on the direction that I have taken and I sincerely thank you for that. Without the support of my family and friends, I would not have been able to undertake such a project. Thank you also to my mother and father who have always believed in me and taught me that with dedication and perseverance you can do whatever you put your mind to!

## Table of Contents

Abstract.....	ii
Preface.....	iii
Acknowledgements.....	iv
Table of Contents.....	vi
List of Tables.....	viii
List of Figures.....	ix
List of Abbreviations.....	x
<b>Chapter 1: Introduction.....</b>	<b>1</b>
1.1 Background.....	1
1.2 Clinical Tests.....	2
1.3 Clinical Decision Guides.....	2
1.4 Summary.....	3
<b>Chapter 2: What is the state of evidence for the assessment of facet joint mediated pain in the cervical spine? A narrative review.....</b>	<b>6</b>
2.1 Background.....	6
2.2 Purpose.....	7
2.3 The cervical facet joint as a source of pain.....	7
2.4 Criterion Standard for the diagnosis of facet joint pain.....	8
2.5 Manual Examination for the identification of facet joint pain.....	10
2.6 Future Implications.....	11
<b>Chapter 3: Intra-rater and Inter-rater reliability of select clinical tests in patients referred for diagnostic facet joint blocks in the cervical spine.....</b>	<b>14</b>
3.1 Introduction.....	14
3.2 Purpose.....	15
3.3 Methods.....	16
3.3.1 Study Design.....	16
3.3.2 Participants.....	16
3.3.3 Procedures.....	17
3.3.4 Data Analysis.....	19
3.4 Results.....	21
3.5 Discussion.....	26
3.5.1 Study Limitations.....	28
3.6 Conclusions.....	29
<b>Chapter 4: Derivation of a positive clinical decision guide of when to refer patients for cervical facet joint blocks.....</b>	<b>30</b>
4.1 Background.....	30
4.2 Purpose.....	31
4.3 Methods.....	32
4.3.1 Study Design.....	32
4.3.2 Participants.....	32
4.3.3 Procedures.....	33
4.3.4 Outcome Measures.....	34
4.3.5 Data Analysis.....	37
4.4 Results.....	39

4.5 Discussion.....	45
4.5.1 Study Limitations.....	49
4.6 Conclusions.....	51
<b>Chapter 5: Derivation of a positive clinical decision guide of when NOT to refer patients for cervical facet joint blocks.....</b>	<b>52</b>
5.1 Background.....	52
5.2 Purpose.....	54
5.3 Methods.....	54
5.3.1 Study Design.....	54
5.3.2 Participants.....	54
5.3.3 Procedures.....	55
5.3.4 Outcome Measures.....	57
5.3.5 Data Analysis .....	58
5.4 Results.....	61
5.5 Discussion.....	69
5.5.1 Study limitations.....	76
5.6 Conclusions.....	78
<b>Chapter 6: Conclusions.....</b>	<b>79</b>
Bibliography.....	84
Appendix A: Self report measures and clinical tests.....	97
Appendix B: Clinical tests.....	105
Appendix C: Bland-Altman plot of the difference between raters for the measure of cervical spine flexion range of motion.....	109
Appendix D: Figure 1: Flow diagram of subject recruitment and participation.....	110
Appendix E: Multivariate logistic regression model with primary exposure variables PST and ER, controlling for age, gender, baseline neck pain and disability, catastrophization, and psychological distress.....	111
Appendix F: ROC curve for the multivariate logistic regression model with primary exposure variables PST and ER, controlling for age, gender, baseline neck pain and disability, catastrophization, and psychological distress.....	112
Appendix G: Multivariate logistic regression model with primary exposure variables MSE and ER, controlling for age, gender, baseline neck pain and disability, catastrophization, and psychological distress.....	113
Appendix H: ROC curve for the multivariate logistic regression model with primary exposure variables MSE and ER, controlling for age, gender, baseline neck pain and disability, catastrophization, and psychological distress.....	114



## List of Tables

Table 3.1: Baseline Demographics.....	22
Table 3.2: Intrarater reliability of cervical range of motion.....	23
Table 3.3: Interrater reliability of cervical range of motion.....	24
Table 3.4: Reliability of symptom response during intrarater and interrater range of motion testing.....	24
Table 3.5: Intrarater and interrater reliability and prevalence of positive ratings for the ER test, MSE, and PST .....	25
Table 4.1: Baseline Demographics.....	40
Table 4.2: Contingency table of frequency of test findings for the manual spinal examination and palpation for segmental tenderness.....	41
Table 4.3: Odds Ratios of the clinical variables in univariate logistic regression for predicting cervical facet joint pain .....	41
Table 4.4: Contingency table of frequency of test findings for the manual spinal examination and the outcome of diagnostic facet joint blocks.....	42
Table 4.5: Contingency table comparing frequency of test findings for the palpation for segmental tenderness and the outcome of diagnostic facet joint blocks.....	42
Table 4.6: Contingency table comparing frequency of test findings for the extension-rotation test and the outcome of diagnostic facet joint blocks .....	42
Table 4.7: Contingency table comparing frequency of test findings for a combination of the MSE and PST and the outcome of diagnostic facet joint blocks.....	43
Table 4.8: Contingency table comparing frequency of test findings for a combination of the MSE, PST, and ER test and the outcome of diagnostic facet joint blocks .....	43
Table 4.9: Accuracy statistics with 95% confidence intervals for the clinical decision guides .....	44
Table 5.1: Baseline Demographics.....	62
Table 5.2: Contingency table of frequency of test findings for the manual spinal examination and palpation for segmental tenderness.....	63
Table 5.3: Odds Ratios of the clinical variables in univariate logistic regression for predicting cervical facet joint pain.....	64
Table 5.4: Contingency table of frequency of test findings for the manual spinal examination and the outcome of diagnostic facet joint blocks.....	64
Table 5.5: Contingency table comparing frequency of test findings for the palpation for segmental tenderness and the outcome of diagnostic facet joint blocks.....	65
Table 5.6: Contingency table comparing frequency of test findings for the extension-rotation test and the outcome of diagnostic facet joint blocks.....	65
Table 5.7: Contingency table comparing frequency of test findings for a combination of the MSE and PST and the outcome of diagnostic facet joint blocks.....	65
Table 5.8: Contingency table comparing frequency of test findings for a combination of the MSE, PST, and ER test and the outcome of diagnostic facet joint blocks.....	66
Table 5.9: Accuracy statistics with 95% confidence intervals for the clinical decision guides.....	67

List of Figures

Figure 1: Flow Diagram of subject recruitment and participation.....110

## List of Abbreviations

CDG – Clinical decision guide  
CPG – Clinical prediction guide  
ER – Extension-rotation test  
GHQ-28 – General Health Questionnaire (28 item)  
ICC - Intraclass correlation coefficient  
 $\kappa$  - kappa  
MSE – Manual spinal examination  
NDI – Neck Disability Index  
NPRS - Numeric Pain Rating Scale  
PCS – Pain Catastrophizing Scale  
PST – Palpation for segmental tenderness  
RFN – Radiofrequency neurotomy  
ROM – Range of motion  
S-LANNS - the self-report version of the Leeds Assessment of Neuropathic Symptoms and Signs pain scale  
95% CI – 95% Confidence Interval

## Chapter 1: Introduction

### 1.1 Background

Neck pain is a common problem and is experienced by up to 23% of adults at any point in time.<sup>1</sup> The 12-month prevalence of neck pain in the general population has been reported to range from 12.15 to 71.5%.<sup>2</sup> The cumulative incidence of self-reported neck pain has been reported to be as high as 179 per 1000 persons.<sup>3</sup> The disability associated with persistent neck pain can have a significant socioeconomic impact, highlighting the public health burden of this type of problem.<sup>4</sup> Individuals with neck pain often seek advice from health care professionals such as medical specialists, primary care physicians, and physiotherapists. In the United States, the annual health care expenditures for neck pain conditions increased by 95% over a 9-year period, with the majority of this increase secondary to costs associated with visits to medical specialists.<sup>4</sup> In many cases, the specific etiology of neck pain presents a challenge to health care professionals. Findings from the patients' history, imaging, and physical examination are not always useful predictors of the cause of their neck pain.

Spinal pain can emanate from the intervertebral disc, nerve roots, facet joint, and ligaments.<sup>5</sup> Studies involving diagnostic facet joint blocks have implicated the facet joint as a primary source of pain in 36% to 67% of individuals with persistent neck pain, suggesting that this joint is a common source of pain.<sup>6-8</sup> Diagnostic facet joint blocks involve the injection of a local anaesthetic onto the sensory nerve from the target facet joint. This occurs over two different occasions, typically one week apart. The objective is to identify the facet joint as a primary source of pain. The

results of this procedure can be used to inform decisions about the appropriateness of radiofrequency neurotomy, which is the intervention typically performed to provide longer-term relief of facet joint mediated pain.<sup>9-12</sup> The costs and wait-times for facet joint interventions in many parts of the world are significant<sup>13</sup>, which highlights the need for a change to the current clinical decision-making and referral process.

## 1.2 Clinical Tests

The accuracy of stand-alone clinical tests to identify patients with primary facet joint pain is questionable. There is evidence to suggest clinicians can identify those with facet joint pain in the neck.<sup>14</sup> Jull et al.<sup>14</sup> reported that a physiotherapist's assessment was 100% sensitive and specific in identifying patients with facet joint pain in the neck. The assessment involved a 'hands-on' evaluation (manual examination) of segmental mobility and pain provocation. Two decades later, these findings were refuted and the validity of the manual examination to identify those with facet joint pain was questioned.<sup>15</sup> In addition, the reliability of the manual examination has reported to be highly variable in patients with mechanical neck pain.<sup>16-21</sup>

## 1.3 Clinical Decision Guides

With the current state of the literature being inconclusive in facilitating clinical decisions on when to refer patients with neck pain for facet joint interventions, to prevent unnecessary procedures and to facilitate appropriate referrals for individuals most likely to respond to these diagnostic injections, further research is needed to identify the optimal clinical tests. The development and use of

a clinical decision guide (CDG) may provide a solution to this problem. Clinical decision guides combine information from the patients' history and physical examination to enhance clinical decision-making in the context of a patients' diagnosis, prognosis, or therapeutic outcome.<sup>22-24</sup> In practice, clinicians incorporate findings from a cluster of tests to facilitate patient management. Seldom does the outcome of one test used in isolation depict the complete clinical picture. Neck pain is a heterogeneous problem. Thus, the use of CDGs to subgroup this patient population to inform optimal intervention pathways is appealing. Clinical decision guides are developed through a rigorous process involving derivation, validation, and impact analysis.<sup>22,25</sup> It has been suggested that clinical tests that possess acceptable levels of reliability can be included in the development of CDGs.<sup>25</sup> Clinical decision guides that incorporate tests that are not reliable are bound for failure in validation research as measurement error and misclassification frequencies will be greater.

#### 1.4 Summary

Previous research in the identification of patients with cervical facet joint pain has provided conflicting results.<sup>14,15</sup> Additionally, this research examined the utility of the manual examination in the diagnosis of facet joint pain, but did not incorporate the findings from a cluster of clinical tests. Thus, previous research has not mirrored the ideal practice patterns of clinicians.<sup>26</sup> Considering the rising costs and significant wait-times associated with facet joint interventions, clinicians could benefit from research outlining the most optimal combination of clinical variables

identifying those that may respond positively to diagnostic facet joint procedures.

Therefore, the following questions will be addressed in this thesis:

- 1) What is the state of evidence for the assessment of facet joint mediated pain in the neck?
- 2) What is the intra-rater and inter-rater reliability of common clinical tests used in patients that have been referred for diagnostic facet joint procedures?
- 3) Can we develop a CDG in the determination of patients that may respond positively to diagnostic facet joint procedures?
- 4) Can we develop a CDG in the determination of patients that may respond negatively to diagnostic facet joint procedures?

For clarification, in Chapters two and three, reference is made to the term clinical prediction guides (CPG) in the context of identifying patients with facet joint pain in the neck. Since the time of preparation and submission of the two manuscripts associated with these chapters, there has been published commentary recommending a change in terminology in diagnosis-based clinical prediction research.<sup>23</sup> The term, “clinical decision guide” has been suggested, as opposed to “clinical prediction rule (CPR) and clinical decision rule (CDR)” as these terms may imply that the “rule” has undergone formal validation, potentially misleading clinicians as to the level of evidence associated with the particular CPR/CDR. The use of the word, “guide” reflects the current state of diagnosis-based research for classifying patients with spinal pain. As this type of research is in its infancy (i.e. further validation and impact analysis needed), a diagnosis-based clinical decision

guide may inform clinicians in their management approaches without negatively impacting patient care by premature widespread use of the CDG.

Chapter two provides a narrative review of the need for evidence-based screening procedures to inform appropriate referral for diagnostic facet joint blocks in the neck. Chapter three is a study that analyzed the reliability of common clinical tests in patients preceding diagnostic facet joint blocks in the neck. The tests that achieved predetermined acceptable levels of reliability were included in the derivation of the clinical decision guide. Chapters four and five report the positive and negative decision guides respectively. Chapter six is the concluding chapter of this thesis and provides a summary of these findings as well as directions for future research.



**Chapter 2: What is the state of evidence for the assessment of facet joint mediated pain in the cervical spine? A narrative review.**  
*Reprinted with permission from Manual Therapy*

2.1 Background

Cervical spine pain is a common condition encountered by physiotherapists and other medical professionals. Based on clinical examination alone, the specific etiology of neck pain can be difficult to diagnose.<sup>27</sup> Nonetheless, in studies involving diagnostic facet joint blocks, the prevalence of facet joint mediated neck pain has been reported to range from 36% to 67%.<sup>6-8</sup>

Patients with persistent, disabling neck pain are increasingly being referred for diagnostic facet blocks to assess their suitability for interventional procedures such as radiofrequency neurotomy (RFN). Utilization of facet joint procedures increased by 624% in the United States between 1997 and 2006.<sup>13</sup> Radiofrequency neurotomy is a minimally invasive neuroablative procedure used to interrupt the pain pathways stemming from the facet joint. Pain medicine experts suggest that individuals may be considered for RFN only if they have experienced a positive response to diagnostic facet joint blocks, as this is an indicator of more substantial treatment benefits from RFN.<sup>10,28,29</sup> There are often lengthy wait-times for diagnostic blocks and RFN. Both incur significant costs to the health care system, which is compounded by referral of patients who respond negatively to the blocks. These issues of patient suitability, wait-time, and costs may be addressed by having screening methods to better identify patients likely to have a positive response to diagnostic facet joint blocks.

Physiotherapists are accessible practitioners within the medical team who have appropriate knowledge and skills in the examination of cervical spine disorders to employ effective screening methods to reduce unnecessary referrals of patients for diagnostic facet joint blocks. However, currently, there is little evidence that any one factor related to patient history, physical examination or medical imaging can predict the outcome of diagnostic facet joint blocks. It is reasoned that the creation and use of a clinical prediction guide (CPG) may foster a more efficient referral system for these diagnostic procedures.

## 2.2 Purpose

The purpose of this professional issues paper is threefold. First, it reviews current concepts of the facet joint as a source of pain, including a description of the criterion standard for the diagnosis of facet joint pain and indications for interventional procedures. Secondly, it explores the utility of manual examination in the clinical diagnosis of cervical facet joint pain. Lastly, it proposes a need for the derivation of a CPG to facilitate patient referral for diagnostic facet joint blocks and presents an overview of a study design to meet this need.

## 2.3 The cervical facet joint as a source of pain

There has been significant laboratory and clinical research implicating the cervical facet joint as a source of pain. Examination of sensory nerve endings in cervical facet joint capsules revealed the presence of free (nociceptive) nerve endings in subsynovial and dense capsular tissues<sup>30</sup> Using immunohistochemistry, Kallakuri et al.<sup>31</sup> identified the presence of nociceptive and inflammatory neuropeptides (substance P and calcitonin gene-related peptide) in human facet

joint capsules.<sup>31-33</sup> In examining conduction velocities, Chen et al.<sup>34</sup> outlined the distribution of A-delta and C-fiber sensory receptors in the facet joint capsule providing further support for the involvement of the facet joint in nociception.<sup>34</sup>

In a clinical study, Dwyer et al.<sup>35</sup> injected contrast medium into facet joints of normal volunteers.<sup>35</sup> From this inaugural work, a composite map was made of the typical patterns of pain referral arising from specific facet joints. Subsequently, this pain map was validated in patients with neck pain who underwent diagnostic facet blocks.<sup>36</sup> As a result, it was suggested that pain maps may be used as part of screening procedures to decipher the facet joint(s) likely responsible for a patient's neck pain.

#### 2.4 Criterion standard for the diagnosis of facet joint pain and interventional pain management

Comparative, controlled medial branch blocks (MBB), under fluoroscopic guidance, have been advocated as the international standard in the diagnosis of facet joint pain.<sup>37,38</sup> The medial branch of the dorsal ramus innervates the facet joint<sup>39</sup> Each facet joint receives sensory supply from the medial branch of the dorsal ramus above and below the joint. Therefore, both sensory nerves need to be anaesthetized in order to block a single facet joint. Medial branch blocks have been shown to be target specific and have been validated against placebo-controlled facet joint blocks.<sup>6,40,41</sup> The procedure involves the injection of a local anaesthetic, on two different occasions (comparative blocks), onto the sensory nerves of the target facet joint. A different anaesthetic (with differing durations of action) is used on each occasion. A positive response is defined as at least 80% relief of familiar pain

intensity following the diagnostic block whose effect is concordant with the duration of the anaesthetic used.<sup>41</sup> Comparative block procedures have been advocated over single block procedures as false positive rates associated with single block procedures range from 27% to 63%.<sup>42,43</sup>

In a recent systematic review, Falco et al.<sup>44</sup> concluded that there was strong support for the effectiveness of RFN in managing chronic and disabling neck pain.<sup>44</sup> In contrast, Niemisto et al.<sup>45</sup> concluded that there was limited evidence that RFN produces short-term pain relief in those with chronic neck pain.<sup>45</sup> The conclusions of the earlier review have been challenged as the authors included studies that used single diagnostic blocks for the diagnosis of facet joint pain, which, as outlined above, possess high false positive rates.<sup>46</sup> Subsequently, patients referred for RFN based on the results of single diagnostic blocks may not be appropriate for the intervention, resulting in negative outcomes. Therefore, it is imperative that the clinical diagnosis be accurate to minimize inappropriate referrals for RFN.

Median duration of substantial pain relief attributed to RFN has been shown to range from 263 – 422 days.<sup>9,10</sup> Although the permanency of pain relief may not seem optimal, patients undergoing this procedure have failed conservative management and continue to have marked limitations in daily activities. Furthermore, the risk associated with diagnostic blocks and RFN are small, and may involve short-term post-operative pain, transient neuritis, and/or dysaesthesias.<sup>28</sup> Importantly, repeat RFN is a viable option for patients who have had success with the initial procedure and have been shown to be successful in reducing pain and disability in the vast majority of patients studied.<sup>10,47,48</sup>

## 2.5 Manual examination for the identification of cervical facet joint pain: Is it a stand-alone test?

Manual examination of the cervical spine provides information regarding the segmental level of dysfunction.<sup>14,16,26</sup> The accuracy and reliability of such examination has been variable.<sup>18,49,50</sup> In the context of the manual examination, segmental paraspinal tenderness has shown to be predictive of positive outcomes following RFN.<sup>51</sup> Jull et al.<sup>14</sup> reported a sensitivity and specificity of 100% for their manual examination in the diagnosis of cervical facet joint pain. The manual examination consisted of a physiotherapist performing a series of procedures involving both passive physiological and passive accessory intervertebral motions of the cervical segments. The physiotherapist evaluated the quality and quantity of segmental motion, as well as pain provocation during the manual examination prior to determining if the facet joint was the primary source of the subject's cervical spine pain. The manual examination was performed as part of a comprehensive subjective and physical evaluation.

These results were challenged 20 years later. King et al.<sup>15</sup> questioned the validity of the diagnostic accuracy statistics presented by Jull et al.<sup>14</sup>, because the criterion standard used to validate the manual examination were single diagnostic facet block procedures, now known to have a high false positive rate. In addition, as confidence intervals were not reported by Jull et al.<sup>14</sup> the point estimates representing sensitivity and specificity of the manual examination may have been misleading.<sup>15</sup> Using comparative, controlled MBB's as the criterion standard, King et al.<sup>15</sup> reported a sensitivity of 89% (95% CI: 82%, 96%), a specificity of 47% (95%

CI: 37%, 57%), and a positive likelihood ratio of 1.7 (95% CI: 1.2, 2.5) for manual examination. They noted that the high sensitivity of manual examination was most likely reflective of the high pre-test probability of facet joint pain in their sample (77%). The authors suggested that the validity of manual examination is reflected by its specificity, and as such, concluded that manual examination of cervical spine facet joint pain lacked validity. One needs to be cautious when interpreting the accuracy of estimates of sensitivity and specificity reported by King et al.<sup>15</sup>.

Measurement bias (verification bias) may partially explain the discrepancy between Jull et al's<sup>14</sup> and King et al's<sup>15</sup> study results, as only the subjects that tested positive on the index test (manual examination) went on to undergo the reference standard comparative, controlled MBB's.<sup>52</sup> This type of bias tends to create an overestimation of sensitivity and an underestimation of specificity.<sup>53</sup>

2.6 Future implications: Is there a need for a clinical prediction guide for those with cervical facet joint pain?

Although the utility of manual examination for determining cervical facet joint pain is promising, there is a need for further research to develop a valid diagnostic paradigm for neck pain, especially if referral for diagnostic blocks is in question. Many clinical tests for neck pain, used in isolation, lack validity for specific anatomical diagnoses.<sup>54</sup> The basis of clinical decision making is formed, in most instances, by the findings of clusters of clinical tests not by any one test alone.<sup>26</sup> A CPG allows the clinician to apply the results of the patient history, self-report measures and physical examination toward a diagnosis, prognosis, or likely response to treatment.<sup>22,52,55</sup> Clinical prediction guides for appropriate referral of

patients for radiography of the ankle and cervical spine post trauma have influenced therapeutic decisions and have been shown to be accurate and cost-effective.<sup>56,57</sup> To date, no such studies have been published regarding the determination of cervical facet joint pain.

The influence of psychological comorbidity on the outcome of facet joint blocks has been reported.<sup>58</sup> Specifically, patients with elevated levels of depression and anxiety may be less likely to experience a positive response to MBB's. Similarly, pain catastrophizing is an important predictor of the pain experience. Therefore, it is valuable to identify patients that may be susceptible to adverse pain responses following invasive medical procedures such as diagnostic facet blocks.<sup>59,60</sup> Consequently, a CPG possibly inclusive of the findings from pain maps, the physical examination, manual examination and psychological screening may decipher those likely to respond to diagnostic facet joint blocks.

To determine if a CPG can be created for identifying those likely responsive to MBB, a study is currently underway in our clinical facility. Utilizing a prospective design, consecutive patients referred to our multidisciplinary spine intervention centre will undergo a comprehensive clinical examination performed by a physiotherapist, including questionnaires to identify any psychological comorbidities, followed by comparative controlled MBB's performed by an interventional radiologist. Multivariable regression analysis will derive the most clinically useful and predictive cluster of variables that physiotherapists may consider when referring patients for diagnostic facet joint blocks. The CPG would then require validation in a prospective study. A CPG may foster optimal

management pathways for patients, and augment efficiency in health care systems by reducing unnecessary referrals for these diagnostic procedures and their associated interventions.



**Chapter 3: Intra-rater and inter-rater reliability of select clinical tests in patients referred for diagnostic facet joint blocks in the cervical spine.**  
*Reprinted with permission from Archives of Physical Medicine and Rehabilitation*

### 3.1 Introduction

Neck pain is common in today's society with an estimated cumulative incidence of 179 cases per 1000 persons.<sup>2</sup> Up to 65% of individuals report neck pain in their lifetime.<sup>1,61</sup> Although specific etiology can be difficult to determine, studies using comparative, controlled facet joint blocks implicate the facet joint as a primary source of pain in 36% to 67% of those with persistent neck pain.<sup>6-8</sup>

The approach for diagnosis of facet joint mediated pain most recognized internationally are controlled block procedures using either two different local anaesthetics or placebo-controlled procedures.<sup>37 62</sup> The use of facet joint procedures in the United States increased by more than 600% between 1997 and 2006.<sup>13</sup> Facet joint blocks are invasive procedures, associated with significant costs and a small element of risk to the patient. There are lengthy wait-times for these procedures in many jurisdictions, where resources are limited. Patients who ultimately respond negatively to diagnostic blocks magnify these wait-times. A clinical method to screen for patients most likely to benefit from diagnostic facet blocks would aid in reducing healthcare costs and wait-times. There is little evidence to suggest that any one factor related to the patients' history or clinical examination can predict the outcome of facet block procedures.<sup>15</sup> Thus, it has been suggested that the derivation of a clinical prediction guide, incorporating findings from a cluster of clinical tests, may provide the clinician with a more accurate

determination of those who may respond positively to diagnostic facet joint blocks.<sup>63</sup>

A clinical test must be deemed reliable before it can be incorporated into a clinical prediction guide.<sup>25,64</sup> Although still controversial, findings from clinical tests such as range of motion, segmental palpation, the extension-rotation test, and manual spinal examination are used as guides to assist clinicians in making management decisions in the context of cervical facet joint mediated pain.<sup>14,51,63</sup> The reliability of the extension-rotation test has not been examined.<sup>51</sup> Studies evaluating the reliability of manual spinal examination (i.e., passive spinal mobility at each segment and pain provocation during segmental motion) in patients with neck pain have reported conflicting results.<sup>16-21</sup> In addition, clear operational definitions of the examination and inclusion of all spinal segments in the cervical spine has been inconsistent.<sup>16,18</sup> No published literature addresses the reliability of these tests in individuals who have been referred for diagnostic facet joint blocks. For the purposes of the present study, these patients may be distinguished from patients with “typical mechanical neck pain” by their reports of persistent symptoms for at least 3 months duration, with higher levels of neck pain and disability, and a failure to respond to conservative rehabilitation and pharmacological interventions.<sup>5</sup>

### 3.2 Purpose

The purpose of this study was to determine the intra-rater and inter-rater reliability of common clinical tests used to evaluate patients with persistent neck pain referred for diagnostic facet joint blocks. Determining the reliability of each

test will enable the derivation of a clinical prediction guide to identify which patients are best suited for diagnostic facet joint blocks.

### 3.3 Methods

#### *3.3.1 Study Design*

This study was a single-group repeated-measures reliability study.

#### *3.3.2 Participants*

Consecutive patients with persistent neck pain, referred to a tertiary interventional pain management centre in Calgary, Alberta, Canada were approached to participate. Subjects were included if aged between 18 and 65 years and reported neck pain intensity of  $\geq 3$  out of 10 on a Numeric Pain Rating Scale (NPRS) for at least the last three months.<sup>65</sup> This standard was set to ensure that a subject's pain intensity exceeded that of the reported measurement error of the NPRS.<sup>66</sup> Subjects were excluded if they presented with: cervical radiculopathy and/or upper motor neuron disease; neck pain related to systemic disease, infection, neoplasm, or fracture; a medically diagnosed psychological disorder; uncontrolled diabetes; uncontrolled clotting disorder; pregnancy; a workers compensation claim or ongoing litigation.

Consecutive sampling methods were applied and of the 108 individuals approached to participate in the study, 27 were excluded (14 were older than 65, 4 possessed a significant language barrier, 3 already had their injection, 1 could not cease their anticoagulant therapy, 4 had pain intensity  $< 3$ , and 1 could not secure transportation to their appointment), 11 declined participation, and 56 consented to participate. There were no clinically relevant differences in age, gender, neck pain

intensity, and duration of neck pain between individuals who participated in the study and those who declined.

At baseline, participants completed a demographic questionnaire, the Neck Disability Index<sup>67</sup>, the Pain Catastrophizing Scale<sup>59</sup>, the General Health Questionnaire – 28<sup>68</sup>, and the self-report version of the Leeds Assessment of Neuropathic Symptoms and Signs pain scale (S-LANSS)<sup>69</sup>. The questionnaires provided background data for the study, and are not reported in the results. Subjects were assessed prior to their first diagnostic facet joint block. Ethical approval for this study was obtained from the Conjoint Health Research Ethics Board at the University of Calgary (Ethics ID: 23783).

### *3.3.3 Procedures*

Once written consent was obtained, subjects completed all questionnaires. They then underwent a standardized clinical examination. The assessors were two experienced physiotherapists with 12-16 years of clinical experience. Select clinical tests commonly used in patients with neck pain were included in the standardized clinical examination. The physiotherapists were provided with a training manual outlining the standardized approach to the clinical examination, including operational definitions of the clinical tests. They underwent a one-hour training session to ensure a standardized approach. Data collection took place between October 2011 and March 2012.

To determine inter-rater reliability, subjects were assessed by the two physiotherapists independently prior to their scheduled diagnostic facet joint blocks. Subjects were given a 5-minute break between testing sessions. Assessor

order was randomized to minimize any potential bias. The second assessor was blinded to the results of the first assessment. Both assessors were blinded to any clinical information pertaining to the subjects (including the level of facet joint block to be performed) to reduce the potential for clinical review bias.<sup>70</sup> Subjects were asked not to reveal any information from the first assessment to the second assessor. Intra-rater reliability was determined by having the same physiotherapist re-examine the same patient seven days following the initial examination.

The clinical examination was performed in the following sequence and included: the assessment of cervical range of motion (ROM), extension-rotation test, manual spinal examination, and palpation for segmental tenderness. Each testing session lasted approximately 15 minutes.

Cervical spine ROM: Measurements of cervical ROM into flexion, extension, left and right side flexion were made with a universal inclinometer<sup>19</sup>. Cervical rotation was measured via a standard dual-armed goniometer.<sup>19</sup> Subjects were seated and asked to maintain a neutral spine position throughout the examination to ensure that thoracic movement was minimized. They were asked to move their head and neck as far as possible in the direction being measured and one measurement was taken for each direction. Subjects were asked to report any pain response on an NPRS, which was subsequently categorized as increased, decreased or no change in familiar pain. Refer to Appendix B for further details.

Extension-rotation test: Subjects were seated and asked to extend their head and neck as far as possible. Rotation was added and subjects reported any pain at the end of motion. A positive test was defined as the provocation of familiar cervical

spine pain intensity ( $\geq 3/10$ ). Rotation was tested to both sides. Refer to Appendix B for further details.

Manual spinal examination: The manual examination was performed to detect the presence or absence of cervical facet joint dysfunction.<sup>14,71</sup> The subject was positioned prone with their cervical spine in a neutral position. The assessor applied a posterior-anterior directed force over the articular pillars from C2-3 to C6-7 on each side (diagnostic facet joint blocks were not performed for the C1-2 and C0-1 joints). Any perceived resistance to motion was categorized as normal, slight, moderate, or marked.<sup>71</sup> The subject reported any pain provocation on a NPRS. The test was considered positive if the subject rated any familiar local or referred pain provoked by the test as  $\geq 3/10$  and the assessor rated moderate or marked resistance to motion.<sup>71</sup> Refer to Appendix B for further details.

Palpation for segmental tenderness: This palpation test is based on the neuroanatomical relationship between the deep cervical segmental muscles and the facet joint capsule.<sup>39,72</sup> With the subject prone, the assessor palpated the segmental muscles overlying the facet joints, C2-3 to C6-7 bilaterally. The test was considered positive if the patient reported an increase in familiar pain, either local or referred, at an intensity of  $\geq 3/10$ . Refer to Appendix B for further details.

#### *3.3.4 Data Analysis*

The sample size was determined based on the primary clinical outcome of interest, which was the segmental tenderness test (with a dichotomous outcome). This test was chosen as it has been suggested as a predictor of a positive outcome following radiofrequency neurotomy.<sup>51</sup> Sim and Wright<sup>73</sup> published a sample size

table based on the goodness of fit formula provided by Donner and Eliasziw<sup>74</sup>. From the table, with the inclusion of a two-tailed test with a null value of kappa of 0.60, at a significance level of  $\leq 0.05$ , and power of 0.8, the sample size required was 56.

Descriptive statistics were used to summarize the subjects' baseline demographic data. The Cohen kappa ( $\kappa$ ) was used to calculate intra- and inter-rater reliability of the clinical tests with two possible response options (extension-rotation test, manual spinal examination, and palpation for segmental tenderness).<sup>75</sup> A kappa value of  $\geq 0.60$  was considered acceptable as per recommended methodological standards in the development of clinical decision rules.<sup>25</sup> Kappa values less than 0.10 indicate no agreement; 0.11 – 0.20 indicate slight agreement; 0.21 – 0.40 indicate fair agreement; 0.41 – 0.60 indicate moderate agreement; 0.61 – 0.80 indicate substantial agreement; and values greater than 0.81 indicated almost perfect or excellent agreement.<sup>76</sup> When judging the magnitude of the kappa coefficients, examination of cross-tabulations allowed us to assess the effects of prevalence and systematic bias on these estimates.<sup>73</sup>

Intraclass correlation coefficients (ICC) model 2,1 (two-way random effects model), and 95% confidence intervals were calculated to determine the inter-rater reliability for cervical ROM. ICC model 3,1 (two-way mixed model), and 95% confidence intervals were calculated to determine intra-rater reliability for cervical ROM. ICC values above 0.75 are indicative of good reliability.<sup>77</sup> Bland-Altman plots were used to provide a visual representation of the differences between raters plotted against the mean score for each subject.<sup>78</sup> The graphic representation allows for assessment of a biased pattern in the scoring between raters.<sup>78</sup> The 95% limits of

agreement were estimated to assess the spread of difference scores between the two raters.

A weighted  $\kappa$  was used to calculate the reliability of categorical data with more than two response options, namely the intra-and-inter-rater reliability for symptom response during cervical ROM testing. The weighted  $\kappa$  was calculated via the linear weighting method.

As a measure of response stability, the standard error of measurement (SEM) was calculated for cervical spine ROM. All statistical analyses were performed using STATA 11 (Texas, USA) statistical software.

### 3.4 Results

The baseline demographic data for the subject group are presented in Table 3.1. Intra- and inter-rater reliability was high for cervical ROM (ICCs ranging from 0.90 to 0.97 for the six directions of motion). Tables 3.2 and 3.3 provide the estimated ICCs (95% confidence intervals), the 95% limits of agreement, and the SEM associated with the cervical ROM measures. Visual inspection of the Bland and Altman plots indicated that agreement was independent of the magnitude of the measurement (See Appendix C for an example related to cervical flexion ROM). For symptom response to ROM testing, the weighted kappa coefficients ranged from 0.60 – 1.0 for intra-and-inter-rater reliability (Table 3.4). Of note, the symptom response, “decreased symptoms” did not occur, forcing the empty cells for observed and expected values for this category to be zero. Coefficients for weighted kappa and kappa would be the same in this instance.



For the extension rotation test, kappa values ranged from 0.72 to 0.75 for intra-rater and 0.92 to 0.93 for inter-rater reliability. For the manual spinal

**Table 3.1: Baseline Demographics**

	<b>Median (Range)</b> <b>N = 56</b>
Age (years)	46 (21 - 64)
Gender	37 females, 19 males
Baseline neck pain intensity (NPRS 0-10)	5 (3 - 9)
Subjects (%) with baseline neck pain intensity $\geq 5/10$	79
Duration of neck pain (months)	13 (3 - 216)
Neck Disability Index (0-50)	19 (8 - 38)
Subjects (%) with Neck Disability Index scores $\geq 15/50$ (self-reported moderate-severe neck pain and disability)	71
Pain Catastrophizing Scale (0-52)	15 (0 - 44)
General Health Questionnaire – 28 (0-84)	21 (9 - 65)
s-LANNS (0-12)	5 (0 - 19)

examination, kappa values ranged from 0.62 to 0.88 for intra-rater and 0.79 to 0.96 for inter-rater reliability. For palpation for segmental tenderness, kappa values ranged from 0.51 to 0.84 for intra-rater and 0.74 to 0.96 for inter-rater reliability. The estimates of kappa and the 95% confidence intervals for intra-rater and inter-rater reliability for these tests are presented in Table 3.5. The prevalence of positive ratings (between assessors) for the manual examination and palpation of segmental

tenderness was greatest at C4-5 and C5-6 segments (Table 3.5). Many subjects presented with positive findings at more than one spinal level. Analysis of cross tabulations did not reveal evidence of bias pertaining to the extent that the assessors disagreed on the proportion of positive or negative findings.

**Table 3.2: Intrarater reliability of cervical range of motion**

<b>Motion</b>	<b>Mean (<math>\pm</math>SD) (degrees)</b>	<b>ICC<sub>3,1</sub> (95% CI)</b>	<b>SEM (degrees)</b>	<b>95% LOA</b>
<b>Flexion</b>	43 (10)	0.94 (0.90 – 0.96)	2.6	-11.9 – 9.9
<b>Extension</b>	43 (13)	0.95 (0.91 – 0.97)	2.9	-11.3 – 9.4
<b>Right side flexion</b>	33 (9)	0.91 (0.85 – 0.95)	2.6	-9.8 – 7.4
<b>Left side flexion</b>	33 (9)	0.93 (0.88 – 0.96)	2.4	-11.7 – 10.0
<b>Right rotation</b>	59 (14)	0.95 (0.92 – 0.97)	3.2	-12.1 – 10.4
<b>Left rotation</b>	57 (14)	0.97 (0.95 – 0.98)	2.5	-12.1 – 10.4

Abbreviations: SD = standard deviation, SEM = standard error of the measurement, LOA = limits of agreement, CI = confidence interval

**Table 3.3: Interrater reliability of cervical spine range of motion**

<b>Motion</b>	<b>Mean (<math>\pm</math>SD) (degrees)</b>	<b>ICC<sub>2,1</sub> (95% CI)</b>	<b>SEM (degrees)</b>	<b>95% LOA</b>
<b>Flexion</b>	43 (12)	0.93 (0.86 – 0.96)	3.4	-10.9 – 6.9
<b>Extension</b>	42 (13)	0.95 (0.91 – 0.97)	3.1	-9.4 – 8.7
<b>Right side flexion</b>	33 (10)	0.90 (0.83 – 0.94)	3.3	-10.4 – 7.5
<b>Left side flexion</b>	33 (10)	0.90 (0.68 – 0.96)	3.2	-10.4 – 4.8
<b>Right rotation</b>	57 (14)	0.95 (0.91 – 0.97)	3.3	-7.7 – 10.6
<b>Left rotation</b>	54 (15)	0.95 (0.71 – 0.98)	3.4	-4.2 – 11.1

Abbreviations: SD = standard deviation, SEM = standard error of the measurement, LOA = limits of agreement, CI = confidence interval

**Table 3.4: Reliability of symptom response during intrarater and interrater range of motion testing**

<b>Motion</b>	<b>Reliability Type</b>	<b>Agreement (%)</b>	<b>Weighted Kappa (95% CI)</b>
Flexion	IntraR	87.5	0.73 (0.18 – 0.86)
	InterR	92.9	0.85 (0.33 – 0.94)
Extension	IntraR	94.6	0.89 (0.40 – 0.96)
	InterR	100	1.0 (0.55 – 1.0)
Right side flexion	IntraR	83.9	0.64 (0.08 – 0.80)
	InterR	100	1.0 (0.53 – 1.0)
Left side flexion	IntraR	80.4	0.60 (0.05 – 0.76)
	InterR	85.7	0.71 (0.17 – 0.85)
Right rotation	IntraR	85.7	0.66 (0.10 – 0.82)
	InterR	91.1	0.80 (0.28 – 0.92)
Left rotation	IntraR	87.5	0.75 (0.23 – 0.88)
	InterR	92.9	0.86 (0.35 – 0.94)

Abbreviations: IntraR = intrarater reliability, InterR = interrater reliability, CI = confidence interval

**Table 3.5: Intrarater and interrater reliability and prevalence of positive ratings for the ER test, MSE, and PST**

Clinical Test	Spinal Level	Agreement (%)		Kappa (95% CI)		Prevalence (%)	
		Right	Left	Right	Left	Right	Left
<b>Intrarater Reliability</b>							
<b>ER</b>	N/A	87.5	87.5	0.73 (0.18 – 0.86)	0.75 (0.23 – 0.88)	29	46
<b>MSE</b>	C2-3	94.6	92.9	0.85 (0.27 – 0.95)	0.63 (-0.04 – 0.85)	20	7
	C3-4	91.1	94.6	0.71 (.084 – 0.87)	0.84 (0.24 – 0.94)	14	18
	C4-5	92.9	94.6	0.82 (0.26 – 0.93)	0.88 (0.36 – 0.96)	23	30
	C5-6	94.6	94.6	0.87 (0.32 – 0.95)	0.88 (0.36 – 0.96)	25	30
	C6-7	89.3	92.9	0.70 (0.11 – 0.86)	0.84 (0.30 – 0.94)	18	29
<b>PST</b>	C2-3	92.9	89.3	0.80 (0.22 – 0.92)	0.51 (-0.06 – 0.76)	21	11
	C3-4	92.9	94.6	0.76 (0.13 – 0.90)	0.82 (0.21 – 0.94)	18	16
	C4-5	89.3	85.7	0.74 (0.18 – 0.88)	0.69 (0.15 – 0.84)	23	30
	C5-6	89.3	89.3	0.73 (0.16 – 0.87)	0.77 (0.23 – 0.89)	25	32
	C6-7	89.3	92.9	0.70 (0.11 – 0.86)	0.84 (0.30 – 0.94)	20	27
<b>Interrater Reliability</b>							
<b>ER</b>	N/A	96.4	96.4	0.92 (0.43 – 0.98)	0.93 (0.45 – 0.98)	36	45
<b>MSE</b>	C2-3	96.4	96.4	0.88 (0.26 – 0.97)	0.81 (0.06 – 0.95)	16	9
	C3-4	92.9	98.2	0.79 (0.19 – 0.92)	0.95 (0.38 – 0.99)	18	20
	C4-5	98.2	98.2	0.96 (0.44 – 0.99)	0.96 (0.45 – 0.99)	27	29
	C5-6	94.6	98.2	0.87 (0.32 – 0.95)	0.96 (0.47 – 0.99)	25	32
	C6-7	96.4	94.6	0.90 (0.34 – 0.97)	0.88 (0.35 – 0.96)	21	29
<b>PST</b>	C2-3	98.2	96.4	0.95 (0.38 – 0.99)	0.86 (0.18 – 0.96)	20	13
	C3-4	94.6	92.9	0.82 (0.21 – 0.94)	0.74 (0.08 – 0.90)	16	13
	C4-5	98.2	92.9	0.96 (0.45 – 0.99)	0.84 (0.30 – 0.94)	29	29
	C5-6	94.6	91.1	0.87 (0.32 – 0.95)	0.81 (0.29 – 0.92)	25	34
	C6-7	91.1	89.3	0.76 (0.18 – 0.89)	0.76 (0.22 – 0.89)	20	29

Abbreviations: ER = extension-rotation test, MSE = manual spinal examination, PST = palpation for segmental tenderness

### 3.5 Discussion

This study is the first study to demonstrate that clinical tests commonly used for patients with persistent neck pain referred for diagnostic facet joint blocks possess moderate to excellent intra-rater and inter-rater reliability as defined by accepted criteria.<sup>76</sup> Clinical tests that possess acceptable inter-rater reliability and thus reflect minimal variability between clinicians, facilitate valid clinical decision making.<sup>79</sup>

Measurement of cervical ROM yielded ICCs for both intra-rater and inter-rater reliability that were reflective of substantial agreement<sup>64</sup>. The lower bounds of the ICC confidence intervals suggest that measurement of cervical spine ROM with a universal inclinometer is appropriate for clinical use. Our ICCs were higher than other reliability studies of neck pain patients, although the 95% confidence intervals overlapped with other studies.<sup>17,19,80</sup> Measurement error was also lower than previously found, but the 95% limits of agreement associated with cervical ROM measurement were similar in magnitude to that of another study of neck pain patients.<sup>19</sup> The rating of symptom response to ROM testing demonstrated moderate to excellent agreement in both intra-rater and inter-rater trials, consistent with previously published research.<sup>17</sup>

This study is also the first to evaluate the reliability of the extension-rotation test in patients with neck pain. The extension rotation test showed substantial to excellent intra-rater and inter-rater reliability respectively.

Manual spinal examination (C2-7) revealed substantial to excellent intra-rater and inter-rater reliability. Higher kappa values were found compared to those

reported in previous studies.<sup>16-19,50,81</sup> Although it is difficult to compare between studies, this improvement may relate to the operational definition and standardized criteria for determining a positive test used in our study. Criteria for a positive test involved a combination of at least moderate motion dysfunction and familiar pain provocation of  $\geq 3/10$ , thereby reducing the ambiguity in determining the test outcome and facilitating consistency between measures. Studies by Jull et al<sup>21</sup> and Hall et al<sup>20</sup>, utilizing a combination of motion evaluation and pain provocation, reported similar kappa values for their manual examination for detecting painful upper cervical joint (C0-C3/4) dysfunction. Lower kappa statistics reported in previous research may stem from the inclusion of asymptomatic subjects.<sup>50</sup> In the study by Fjellner et al<sup>50</sup>, the mean prevalence of positive findings was 11% or less at many cervical levels. Use of a homogenous, asymptomatic sample may have lead to artificially lower estimates of reliability due to the effects of prevalence bias.<sup>82</sup>

Similarly to what was observed for the manual spinal examination, moderate to excellent intra-rater and inter-rater reliability was found for palpation for segmental tenderness. The higher reliability ratings in this study contrasted the results of other studies<sup>81,83-85</sup>, which neither standardized the test procedure nor outcome to include both familiar pain provocation and a minimum cut-score for pain intensity. Our findings were similar to those reported by Hubka et al<sup>86</sup> whose estimates of inter-rater reliability (kappa = 0.68) were closer to those reported in our study. The investigators determined the test to be positive based on the patient's indication of which spinal segment was the most tender. As in our study,

the addition of specific criteria pertaining to pain intensity may have facilitated greater agreement between clinicians.

In our study, the estimated kappa values associated with the reliability of the extension-rotation test, manual spinal examination, and palpation for segmental tenderness reflected slight superiority for ratings of inter-rater reliability. Similar findings were noted for weighted kappa values related to clinician rating of symptom response to ROM testing. Although it is not clear why this occurred, the evaluation of the confidence intervals associated with the estimated kappa and weighted kappa values are overlapping suggesting they may not be truly different. Since the intra-rater reliability was based on tests that were repeated a week apart, we speculate that this result could reflect a true variability in the severity of the patients reported symptoms as much as an inherent measurement error of the specific test. Of relevance, the clinical tests possessed moderate to excellent agreement for both intra-rater and inter-rater reliability.

### *3.5.1 Study Limitations*

There are potential limitations in this study. Caution is needed in the interpretation of the precision of estimates of kappa for the extension rotation, manual segmental examination, and palpation for segmental tenderness tests, as the associated confidence intervals were wide. This may be reflective of the sample size and the relatively low prevalence of positive findings at each spinal segment. Our assessors were experienced physiotherapists, thus the findings may not necessarily be generalizable to all clinicians. However, the comprehensive, time efficient, and

standardized approach to the clinical tests utilized in this study fosters ease of use in clinical practice.

Our study sample of individuals with persistent neck pain referred for diagnostic facet joint blocks may comprise a homogeneous, more severe group among the spectrum of patients with neck pain. Reliability coefficients associated with clinical tests used in homogeneous samples may be artificially deflated due to low between subject variability (compared with that of error variance) and the increased probability of chance agreement.<sup>77</sup> While our sample involved subjects with neck pain of longer duration, and higher self-reported neck pain and disability than other reliability studies of neck pain<sup>17,18</sup>, it is representative of typical patients referred for diagnostic facet joint procedures.<sup>5</sup> Furthermore, the demographics of our sample revealed heterogeneity in neck pain intensity, duration of symptoms, and disability (Table 1). As a result, we believe that our estimated reliability coefficients were not biased in this regard.

### 3.6 Conclusion

The standardized clinical tests used in our study to assess subjects with neck pain demonstrated clinically acceptable intra-rater and inter-rater reliability. Further research to evaluate the clinical utility of the tests, as part of a clinical prediction guide to identify candidates for diagnostic facet block procedures, is warranted and is currently underway.



## **Chapter 4: Derivation of a positive clinical decision guide of when to refer patients for cervical facet joint blocks**

### 4.1 Background

Neck pain is a common clinical condition encountered by medical and allied health care professionals. The etiology of neck pain can be multifactorial making it a challenge to diagnose and manage. In addition, neck pain is typically episodic in nature, affecting up to 65% of individuals in their lifetime.<sup>1,61</sup> In the United States, self-reported spine pain and associated medical expenditures have increased in recent years, corresponding with a reduction in population health status.<sup>87</sup> This finding highlights the socioeconomic impact of spine pain in the general population. Neck pain can stem from various structures including the intervertebral disc, nerve root, facet, and ligaments.<sup>88</sup> Although specific etiology can be difficult to determine, studies using comparative, controlled facet joint blocks implicate the facet joint as a primary source of pain in 36% to 67% of those with persistent neck pain.<sup>6-8</sup> Patients undergoing facet joint interventions may be distinguished from patients with “typical mechanical neck pain” by their reports of persistent symptoms for at least three months duration, with higher levels of neck pain and disability, and a failure to respond to conservative rehabilitation and pharmacological interventions.<sup>5</sup>

The approach for diagnosis of facet joint mediated pain most recognized internationally are controlled block procedures using either two different local anaesthetics or placebo-controlled procedures.<sup>37 62</sup> The use of facet joint procedures in the United States increased by more than 600% between 1997 and 2006.<sup>13</sup> Facet joint blocks are invasive procedures, associated with significant costs and a small element of risk to the patient. There are lengthy wait-times for these

procedures in many jurisdictions where resources are limited. Patients who ultimately respond negatively to diagnostic blocks magnify these wait-times. A clinical method to screen for patients most likely to benefit from diagnostic facet blocks would aid in reducing healthcare costs and wait-times. There is little evidence to suggest that any one factor related to the patients' history or clinical examination can predict the outcome of facet block procedures.<sup>15</sup> Thus, it has been suggested that the derivation of a clinical decision guide, incorporating findings from a cluster of clinical tests, may provide the clinician with a more accurate determination of those who may respond positively to diagnostic facet joint blocks.<sup>63</sup>

Although still controversial, findings from clinical tests such as range of motion, segmental palpation, the extension-rotation test, and manual spinal examination are used as guides to assist clinicians in making management decisions in the context of cervical facet joint mediated pain.<sup>14,51,63</sup> When examining the accuracy of clinical tests in diagnosing facet joint pain, one needs to consider tests that possess target specificity in identifying a particular spinal level as a potential pain generator. This is necessary since diagnostic facet joint block procedures target specific facet joints and are not applied generally to all facet joints in the neck for each patient.

#### 4.2 Purpose

The purpose of this study was to derive a clinical decision guide (CDG) to identify which patients are best suited for diagnostic facet joint blocks. Specifically, we aimed to answer the question: Are the findings from the manual spinal

examination, palpation for segmental tenderness, and the extension-rotation test predictive of the outcome of diagnostic facet joint blocks in the cervical spine?

### 4.3 Methods

#### 4.3.1 *Study Design*

This study was a prospective cohort study conducted in a tertiary interventional pain management centre in Calgary, Alberta, Canada.

#### 4.3.2 *Participants*

Consecutive patients with persistent neck pain, referred for diagnostic facet joint blocks were approached to participate. Subjects were included if aged between 18 and 65 years and reported neck pain intensity of  $\geq 3$  out of 10 on a Numeric Pain Rating Scale (NPRS) for at least the last three months.<sup>65</sup> This standard was set to ensure that a subject's pain intensity exceeded that of the reported measurement error of the NPRS.<sup>66</sup> Subjects were excluded if they presented with: cervical radiculopathy and/or upper motor neuron disease; neck pain related to systemic disease, infection, neoplasm, or fracture; a medically diagnosed psychological disorder; uncontrolled diabetes; uncontrolled clotting disorder; pregnancy; a workers compensation claim or ongoing litigation.

Consecutive sampling methods were applied and of the 177 individuals approached to participate in the study, 38 were excluded based on the inclusion and exclusion criteria, 14 declined participation, and 125 consented to participate. Refer to Figure 1 (Appendix D) for a description of subject recruitment and participation. There were no clinically relevant differences in age, gender, neck pain intensity, and duration of neck pain between individuals who participated in the study and those

who declined. Ethical approval for this study was obtained from the Conjoint Health Research Ethics Board at the University of Calgary (Ethics ID: 23783).

#### *4.3.3 Procedures*

Once written consent was obtained, subjects completed a baseline demographic questionnaire followed by four standardized self-report measures. The Neck Disability Index (NDI) was completed to assess self-reported neck pain and disability.<sup>67</sup> Higher scores on the NDI reflect greater self-reported disability related to neck pain. The NDI has been shown to have a high degree of test-retest reliability (ICC: 0.89-0.94) and internal consistency (coefficient alpha values: 0.80-0.87).<sup>67</sup> The NDI has been reported to be a valid and responsive measure in those with mechanical neck disorders.<sup>66,89,90</sup> The Pain Catastrophizing Scale (PCS) was used to evaluate catastrophic thoughts in response to the patients neck pain.<sup>59</sup> From a clinical perspective, the PCS may be a useful tool in identifying patients that may be susceptible to adverse pain responses following invasive medical procedures. The PCS has demonstrated excellent internal consistency and moderate to substantial test-retest reliability.<sup>59,60</sup> The General Health Questionnaire – 28 (GHQ-28) is a reliable and valid measure of psychological distress in medical settings.<sup>68,91</sup> The self-report version of the Leeds Assessment of Neuropathic Symptoms and Signs pain scale (S-LANSS) is a measure identifying pain of predominantly neuropathic origin.<sup>69</sup> The S-LANNS has demonstrated construct validity and excellent test-retest reliability.<sup>69,92</sup> Refer to Appendix A for a description of the self-report measures used in this study.

Subjects were assessed prior to their first diagnostic facet joint block. The subjects underwent a standardized clinical examination, which was performed by four experienced physiotherapists with 10-25 years of clinical experience. Select clinical tests commonly used in patients with neck pain were included in the standardized clinical examination. The physiotherapists were provided with a training manual outlining the standardized approach to the clinical examination, including operational definitions of the clinical tests. They underwent a one-hour training session to ensure a standardized approach. The physiotherapists were blinded to the level of the facet joint block to be performed for each subject. Data collection took place between October 2011 and December 2012.

The clinical examination was performed in the following sequence and included: the assessment of cervical range of motion (ROM), extension-rotation test, manual spinal examination, and palpation for segmental tenderness. Each testing session lasted approximately 15 minutes. The standardized clinical tests have demonstrated moderate to excellent intra-rater and inter-rater reliability in patients with axial neck pain referred for diagnostic facet joint blocks.<sup>93</sup>

#### *4.3.4 Outcome Measures*

##### Cervical spine ROM:

Measurements of cervical ROM into flexion, extension, left and right side flexion were made with a universal inclinometer<sup>19</sup>. Cervical rotation was measured via a standard dual-armed goniometer.<sup>19</sup> Subjects were seated and asked to maintain a neutral spine position throughout the examination to ensure that thoracic movement was minimized. They were asked to move their head and neck

as far as possible in the direction being measured and one measurement was taken for each direction. Subjects were asked to report any pain response on an NPRS, which was subsequently categorized as increased, decreased or no change in familiar pain. Refer to Appendix B for further details.

Extension-rotation test:

Subjects were seated and asked to extend their head and neck as far as possible. Rotation was added and subjects reported any pain at the end of motion. A positive test was defined as the provocation of familiar cervical spine pain intensity ( $\geq 3/10$ ). Rotation was tested to both sides. Refer to Appendix B for further details.

Manual spinal examination:

As per previous reports, the manual examination was performed to detect the presence or absence of cervical facet joint dysfunction.<sup>14,71</sup> The subject was positioned prone with their cervical spine in a neutral position. The assessor applied a posterior-anterior directed force over the articular pillars from C2-3 to C6-7 on each side (diagnostic facet joint blocks were not performed for the C0-1 and C1-2 joints). Any perceived resistance to motion was categorized as normal, slight, moderate, or marked.<sup>71</sup> The subject reported any pain provocation on a NPRS. The test was considered positive if the subject rated any familiar local or referred pain provoked by the test as  $\geq 3/10$  and the assessor rated moderate or marked resistance to motion.<sup>71</sup> Refer to Appendix B for further details.

Palpation for segmental tenderness:

This palpation test is based on the neuroanatomical relationship between the deep cervical segmental muscles and the facet joint capsule.<sup>39,72</sup> With the subject

prone, the assessor palpated the segmental muscles overlying the facet joints, C2-3 to C6-7 bilaterally. The test was considered positive if the patient reported an increase in familiar pain, either local or referred, at an intensity of  $\geq 3/10$ . Refer to Appendix B for further details.

Criterion standard diagnostic facet joint block:

Comparative medial branch blocks (MBB) were performed as the criterion standard for the diagnosis of facet joint pain.<sup>37</sup> Comparative blocks have been shown to minimize the risk of false positive responses reported by single block procedures.<sup>42</sup> This procedure involved the injection of 0.5 ml of either Bupivacaine 0.5% or Lidocaine 2% under fluoroscopic guidance onto the sensory nerves (medial branch of the dorsal ramus and/or the third occipital nerve) of the target facet joint. Contrast material (0.25 ml of Omnipaque 300) was injected at each spinal level to ensure target specificity of the facet blocks. All subjects underwent the criterion standard. The anaesthetics were delivered in a random order. The subjects and the interventional radiologist or physiatrist performing the injection were blinded to the anaesthetic used. The interventional radiologist, physiatrist, and radiology technician were blinded to any study related pre-injection outcomes measures or clinical examination findings.

All subjects underwent an initial MBB. A radiology technician recorded the subjects' neck pain intensity on an 11-point NPRS before and after the block. A positive response was defined as a  $\geq 80\%$  decrease in familiar neck pain intensity for at least the duration of the anaesthetic used ( $\geq 1$  hour for Lidocaine and 3 hours for Bupivacaine).<sup>41</sup> Positive responders underwent a second MBB one week

following their initial block. A subject was deemed to have facet joint pain if they were defined as a positive responder to both MBB's. If the subject reported < 80% relief in familiar neck pain intensity following the MBB, they were defined as a negative responder. In the negative responders, if the interventional radiologist or physiatrist performing the injection, or the referring physician felt that another facet joint is likely the putative source of neck pain, the subject may have undergone another facet joint block at a neighboring spinal level.

#### 4.3.5 Data Analysis

The sample size was determined *a priori* based on the reported prevalence (36%-67%) of facet joint pain in the cervical spine.<sup>6-8</sup> With conservative use of this data, we estimated the prevalence of facet joint pain in our sample to be 40%. In deriving a CDG from multivariable regression analyses, it has been stated that at least 10 outcome events (diagnosis of facet joint pain) should occur for each predictor variable.<sup>22,94</sup> *A priori*, we determined that a CDG with more than five predictor variables may not be efficient for clinicians utilizing the guide in practice. From this, our study would require at least 50 positive outcome events (subjects diagnosed with facet joint pain). Given an anticipated prevalence of facet joint pain in our sample of 40% and the use of up to five predictor variables in the CDG, the number of subjects needed in our study was 125 ( $50/n = .40$ ).

Descriptive statistics were used to summarize the subjects' baseline demographic data. Clinical tests and/or demographic variables previously shown in the literature to be associated with the outcome of diagnostic cervical facet joint blocks were analyzed for their association with the outcome of diagnostic facet



blocks in our study via univariate and multivariate logistic regression. Clinical tests with acceptable levels of intra-rater and inter-rater reliability (kappa statistic  $\geq 0.60$  or intraclass correlation coefficient  $\geq 0.80$ ) were considered in the development of the CDG.<sup>25</sup> For the regression analyses, the beta coefficients (and their standard errors) in the various models were evaluated for the effects of collinearity.<sup>95</sup> In addition, contingency tables were constructed for these variables to evaluate the extent of their agreement. If collinearity was detected then the data from clinically relevant variables and the diagnostic outcome of the facet joint blocks was entered into contingency tables and the sensitivity, specificity, and likelihood ratios of the CDG's were calculated, along with their 95% confidence intervals. The objective was to develop a CDG that possessed the highest possible specificity and positive likelihood ratio in determining those with facet joint pain, while maximizing sensitivity. Interpretation of the magnitude of the likelihood ratios followed the guide reported by Guyatt and colleagues<sup>52</sup>, whereby likelihood ratios greater than 10 or less than 0.1 infer large and often conclusive shifts from pre-test to post-test probability; likelihood ratios from five to 10 and 0.1 to 0.2 infer a moderate shift in probability; likelihood ratios from two to five and 0.2 to 0.5 infer small, but sometimes important changes in probability; and likelihood ratios from one to two and 0.5 to 1 infer a change in probability that is rarely useful. All statistical analyses were performed using STATA 11 (StataCorp LP, College Station, Texas, USA) statistical software.

#### 4.4 Results

The baseline demographic data for the subjects, including their median scores (range) on the self-report measures, are presented in Table 4.1. The C5-6, C6-7, and C2-3 facet joints were the most frequent joints to undergo diagnostic facet joint blocks, with a prevalence of 36%, 33%, and 23% respectively. Of the 125 subjects, 52 had positive responses to comparative medial branch blocks, which amounts to a prevalence (pre-test probability) of facet joint pain of 42% in this sample. Of those that were positive responders, 14 were positive at C2-3; 12 were positive at C3-4; four were positive at C4-5; 21 were positive at C5-6; and 11 were positive at C6-7. Of these patients, 10 were positive at two levels (ie; C2-3 and C3-4).

In consideration of the target specificity needed to determine the appropriate facet joints to undergo diagnostic injections, the manual spinal examination and palpation for segmental tenderness tests were selected for regression analysis and the derivation of the CDG. The extension-rotation test was also selected as it has been examined as a potential test to identify those with facet joint pain.<sup>51</sup>

In analyzing the results of the multivariate logistic regression analysis, collinearity was evident in the model with the variables representing the manual spinal examination and palpation for segmental tenderness due to the high agreement between their findings (Table 4.2). For this reason, univariate logistic regression was utilized to describe the association between the manual spinal examination, palpation for segmental tenderness, and the extension-rotation test and the outcome of diagnostic facet joint blocks. In addition, contingency tables were generated and diagnostic accuracy statistics were calculated for these tests.

**Table 4.1: Baseline Demographics**

	<b>Median (Range)</b> <b>N = 125</b>
Age (years)	49 (21 - 65)
Gender	84 females, 41 males
Baseline neck pain intensity (NPRS 0-10)	6 (3 - 9)
Subjects (%) with baseline neck pain intensity $\geq 5/10$	82
Onset of pain (%)	
Traumatic	47
Gradual	46
Sudden	7
Motor vehicle collision (%) (n=53)	
Frontal impact	15
Side impact	26
Rear impact	59
Smoker (%)	22
Employed (%)	
Currently employed	63
On leave due to neck problem	10
Unemployed	11
Retired	11
Student	5
Duration of neck pain (months)	18 (3 - 216)
Neck Disability Index (0-50)	20 (3 - 38)
Subjects (%) with Neck Disability Index scores $\geq 15/50$ (self-reported moderate-severe neck pain and disability)	74
Pain Catastrophizing Scale (0-52)	13 (0 - 44)
General Health Questionnaire - 28 (0-84)	22 (8 - 65)
s-LANNS (0-12)	8 (0 - 19)

From the univariate logistic regression analysis, the odds ratio associated with the findings of the manual spinal examination, palpation for segmental tenderness, and extension-rotation test was 29.71 (95% CI: 9.51 – 92.81), 43.28 (95% CI: 12.11 – 154.77), and 6.85 (95% CI: 2.91 – 16.13) respectively (Table 4.3).

**Table 4.2: Contingency table of frequency of test findings for the manual spinal examination and palpation for segmental tenderness**

MSE	PST		Total
	Positive	Negative	
Positive	66*	3	69
Negative	3	53*	56
Total	69	56	125

Abbreviations: MSE = manual spinal examination, PST = palpation for segmental tenderness, \* = high levels of agreement between the two clinical tests

**Table 4.3: Odds Ratios of the clinical variables in univariate logistic regression for predicting cervical facet joint pain**

Clinical Variable	Coefficient	Standard Error	Odds Ratio (95% CI)
Extension-rotation test	1.92	0.44	6.85 (2.91 – 16.13)
Manual spinal examination	3.39	0.58	29.71 (9.51 – 92.81)
Palpation for segmental tenderness	3.77	0.65	43.28 (12.11 – 154.77)

Abbreviations: CI = confidence interval

**Table 4.4: Contingency table of frequency of test findings for the manual spinal examination and the outcome of diagnostic facet joint blocks**

MSE	Diagnostic facet joint block		Total
	Positive	Negative	
Positive	48	21	69
Negative	4	52	56
Total	52	73	125

Abbreviations: MSE = manual spinal examination

**Table 4.5: Contingency table comparing frequency of test findings for the palpation for segmental tenderness and the outcome of diagnostic facet joint blocks**

PST	Diagnostic facet joint block		Total
	Positive	Negative	
Positive	49	20	69
Negative	3	53	56
Total	52	73	125

Abbreviations: PST = palpation for segmental tenderness

**Table 4.6: Contingency table comparing frequency of test findings for the extension-rotation test and the outcome of diagnostic facet joint blocks**

ER	Diagnostic facet joint block		Total
	Positive	Negative	
Positive	43	30	73
Negative	9	43	52
Total	52	73	125

Abbreviations: ER = extension-rotation test

**Table 4.7: Contingency table comparing frequency of test findings for a combination of the MSE and PST and the outcome of diagnostic facet joint blocks**

MSE and PST	Diagnostic facet joint block		Total
	Positive	Negative	
Positive	48	18	66
Negative	4	55	59
Total	52	73	125

Abbreviations: MSE = manual spinal examination, PST = palpation for segmental tenderness

For the manual spinal examination and palpation for segmental tenderness, positive findings occurred in 69 subjects (Table 4.4 and Table 4.5). For the extension-rotation test, positive findings occurred in 73 subjects (Table 4.6). Sixty-six subjects tested positive on both the manual spinal examination and palpation for segmental tenderness (Table 4.7), whereas, 53 subjects tested positive on all three clinical tests (Table 4.8).

**Table 4.8: Contingency table comparing frequency of test findings for a combination of the MSE, PST, and ER test and the outcome of diagnostic facet joint blocks**

MSE/PST/ER	Diagnostic facet joint block		Total
	Positive	Negative	
Positive	41	12	53
Negative	11	61	72
Total	52	73	125

Abbreviations: MSE = manual spinal examination, PST = palpation for segmental tenderness, ER = extension-rotation test

If a subject tested positive on both the manual spinal examination and palpation for segmental tenderness, the positive likelihood ratio was 3.74 (95% CI:

**Table 4.9: Accuracy statistics with 95% confidence intervals for the clinical decision guides**

Clinical Decision Guide	Sensitivity* (95% CI)	Specificity* (95% CI)	Positive Likelihood Ratio (95% CI)	Post-test Probability of a Diagnosis of Facet Joint Pain*†‡
CDG 1: MSE	92 (88 – 97)	71 (63 – 79)	3.17 (2.22 – 4.64)	70
CDG 2: 2. PST	94 (90 – 98)	73 (65 – 80)	3.48 (2.35 – 5.03)	72
CDG 3: ER	83 (76 – 89)	59 (50 – 68)	2.02 (1.49 – 2.72)	59
CDG 4: Subjects who tested positive on both MSE and PST	92 (88 – 97)	75 (68 – 83)	3.74 (2.49 – 5.63)	73
CDG 5: Subjects who tested positive on MSE, PST, and ER	79 (72 – 86)	84 (77 – 90)	4.94 (2.80 – 8.20)	78

Abbreviations: CDG = clinical decision guide, MSE = manual spinal examination, PST = palpation for segmental tenderness, ER = extension-rotation test, CI = confidence interval

\* = proportions are stated as a percentage (%)

† = assumes a pre-test probability of 42%

‡ = post-test probability given a positive index test is equivalent to the positive predictive value of the index test or cluster of tests

2.49 – 5.63) and the post-test probability of a diagnosis of facet joint pain increased from 42% to 73%. If the subject tested positive on all three clinical tests, the positive likelihood ratio was 4.94 (95% CI: 2.80 – 8.20) and the post-test probability of a diagnosis of facet joint pain increased from 42% to 78%. Refer to Table 4.9 for the accuracy statistics associated with the CDG's.

#### 4.5 Discussion

This study is the first study to derive a CDG, incorporating the findings from a cluster of clinical tests, with utility to predict a diagnosis of cervical facet joint pain. Our results showed that the manual spinal examination and palpation for segmental tenderness exhibit high sensitivity (92% - 94%), in isolation or in combination (CDG 1, 2, or 4), supporting their use as potential screening tests prior to considering referring patients for facet joint blocks. Specificity (84%) of the CDG is maximized when a patient tests positive on the manual spinal examination, palpation for segmental tenderness, and the extension-rotation test (CDG 5). Coinciding with a tests specificity, the positive likelihood ratio is a useful statistic for quantifying a shift in probability favoring the condition of interest when the clinical test is positive.<sup>96</sup> In the context of our study, a clinical test, or cluster of clinical tests, that possess a high positive likelihood ratio would increase clinician confidence in making the decision to refer a patient for a diagnostic facet joint block. Our data indicated that the highest positive likelihood ratio (4.94) was associated with the cluster of test findings from all three clinical tests (CDG 5). The magnitude of this positive likelihood ratio provides clinicians with a small, and near moderate, shift in



probability that the patient has facet joint pain, given positive test findings on all three clinical tests.<sup>52,97</sup>

The aforementioned clinical tests were included in the derivation of the CDG's as there has been some literature indicating their use in evaluating those with potential cervical facet joint pain.<sup>14,15,51</sup> Furthermore, tests such as the manual spinal examination and palpation for segmental tenderness aim to identify specific dysfunctional spinal motion segments. In clinical practice, diagnostic facet joint blocks target presumed painful facet joints. The aim of these injections is not to target every facet joint in the cervical spine in every patient with persistent neck pain, as this would not be resourceful. Hence, clinical tests that possess target specificity are appropriate for consideration in our CDG aiming to identify those that may respond positively to diagnostic facet joint blocks.

Our study is the first to evaluate the accuracy of the extension-rotation test in the diagnosis of cervical facet joint pain. Our data indicated that positive findings on the extension-rotation test (specificity = 59% and positive likelihood ratio = 2.02), when considered as a stand-alone test, might not substantiate a clinician's confidence that the patient has facet joint mediated pain. In addition, our study results foster a similar conclusion for the findings of the manual spinal examination and palpation for segmental tenderness when used in isolation.

Siegenthaler et al.<sup>98</sup> assessed mechanical pain sensitivity (pressure pain thresholds measured in kPa), via an electronic pressure algometer, in 33 patients with unilateral neck pain. They hypothesized that the pain threshold would be lower over the painful facet joint compared with the contralateral side, attempting to

validate segmental tenderness as clinical test in identifying facet joint pain. Using a cut-point of 1 kPa (mean difference between the painful side and non-painful side) to categorize the pressure algometry as positive or negative and comparing these findings with the outcome of comparative, controlled medial branch blocks, the authors reported a sensitivity of 67%, specificity of 16%, and positive likelihood ratio of 0.79. Using a much higher cut-point of 30 kPa, the tests sensitivity decreased to 13%, specificity increased to 95% and the positive likelihood ratio improved to 2.53. Although the use of dichotomous cut-points can be problematic in prediction research as it has been shown to diminish the discriminative ability of the model or test, the choice of a higher cut-point may differentiate individuals that are ‘truly’ positive from those that are not.<sup>99</sup> In our study, the assessors recorded a positive test for segmental tenderness when the subject reported “familiar pain” along with a cut-off of neck pain intensity of  $\geq 3/10$  on a NPRS. The addition of “familiar pain” provocation in the operational definition of the test, along with the use of the clinician’s hands directly palpating taut bands of segmental muscle (compared to an electronic device) may aid in correctly identifying those with a specific anatomical source of pain. When interpreting the precision of the findings of Siegenthaler et al.<sup>98</sup>, one needs to be cautious as the study was not powered to determine diagnostic accuracy statistics. In consideration of the small sample size, the confidence intervals associated with the accuracy estimates (which the authors did not report) would be wide questioning the precision of estimated sensitivity, specificity, and positive likelihood ratio. In addition, the outcome of pressure pain threshold testing in patients with neck pain is known to be highly variable.<sup>100</sup> Using a mean difference

in kPa to formulate the cut-point for a positive test is problematic when dealing with highly variable outcomes, presenting another reason to be cautious when interpreting the reliability of their accuracy statistics.

A landmark study by Jull et al.<sup>14</sup> examined the ability of an experienced physiotherapist to identify symptomatic facet joints in 20 patients with neck pain undergoing single diagnostic facet joint blocks, which was the criterion standard at the time of the study. The physiotherapist was able to correctly identify those with and without facet joint pain, along with the specific segmental level of pain in all 20 patients. Remarkably, the manual spinal examination used in their study was 100% sensitive and 100% specific in deciphering those with and without facet joint pain. In our study, we incorporated a similar manual spinal examination procedure, but we compared our index test findings against those of the currently accepted criterion standard, comparative, controlled medial branch blocks, for the diagnosis of facet joint pain.<sup>37</sup> In the case of single diagnostic blocks, we now know that they possess a high false positive rate ranging from 27% to 63%.<sup>42,43</sup> This finding undermines the magnitude of the sensitivity and specificity noted in the study by Jull et al.<sup>14</sup>, and may explain some of the discrepancies between our study findings.

Using comparative, controlled medial branch blocks as the criterion standard, King et al.<sup>15</sup> studied the validity of the manual spinal examination in assessing patients with neck pain. The manual spinal examination mimicked the procedure used by Jull et al.<sup>14</sup> The authors reported a sensitivity of 89% (95% CI: 82% - 96%), specificity of 47% (95% CI: 37% - 57%), and a positive likelihood ratio of 1.7 (95% CI: 1.2 - 2.5), questioning the criterion validity of the manual spinal

examination. In their study, King et al.<sup>15</sup> noted that only patients with “clinically positive” joints (secondary to the results of the manual spinal examination) underwent the criterion standard facet joint blocks. If the criterion standard is provided solely to patients that test positive on the index test, measurement bias (verification bias) may affect the accuracy statistics of the index test under study. Verification bias may explain the differences in our study findings compared with that of King et al.<sup>15,52</sup> This form of bias may have lead to an underestimate of the specificity of the manual spinal examination noted in the study by King et al.<sup>15,53</sup> Considering the discrepancies in the accuracy of the manual spinal examination noted in the literature to date, it provides further support for the use of CDG’s that utilize the findings from a cluster of clinical tests, rather than the findings of any one test alone, ultimately mimicking the basis of clinical decision making.<sup>26</sup>

#### *4.5.1 Study Limitations*

There are potential limitations in this study. Our assessors were experienced physiotherapists, thus the findings may not necessarily be generalizable to all clinicians. However, the comprehensive, time efficient, and standardized approach to the clinical tests utilized in this study fosters ease of use in clinical practice.

The subjects underwent diagnostic facet joint blocks to the putatively painful joints based on their pain pattern and response to the previous facet joint block. Since they did not receive facet blocks at every spinal level in the neck, some might suggest that this compromises the validity of our diagnostic accuracy statistics. Although this raises a methodological concern, it is neither realistic nor efficient, from a clinical perspective to put every patient through this number of injections. By

having every subject undergo the criterion standard, we believe that we minimized bias related to diagnostic work-up.<sup>53</sup>

The CDG's in our study were derived from the combination of tests that optimized specificity and the positive likelihood ratio calculated from contingency tables. In recent years, the use of multivariate regression modeling to formulate a CDG has provided a method of predicting an outcome from multiple risk factors while simultaneously evaluating effect modification and controlling for confounding. In our study, we reported the outcome of univariate regression analyses as the outcome of multivariate analyses was deemed problematic due to the effects of collinearity related to the manual spinal examination and palpation for segmental tenderness test. The existence of collinearity may inflate the variances related to the estimated coefficients in a model resulting in a lack of statistical significance of one or more independent variables, although the overall model may be significant.<sup>95</sup> Although one can combine highly correlated variables into a single variable for analysis in a multivariate model, we believe that the constructs underlying the manual spinal examination and palpation for segmental tenderness are distinctive for each test, as the former test involves an assessment of mobility. In our view, this renders the information obtained from each test clinically relevant and unique in its own right. In this context, reporting of univariate regression analyses was informative, while avoiding erroneous reporting of its multivariate counterpart.

This study may have omitted potential predictor variables that could have been incorporated in the derivation of the CDG. We chose to include clinical tests that have either been studied previously for their diagnostic accuracy in those with

cervical facet joint pain or have been widely used by clinician's based on the assumption of the ability of the test to provoke facet joint pain. We avoided including multiple predictor variables that did not have a biologically plausible association with facet joint pain, to minimize the risk of analyzing multiple variables without enough power to detect important predictors.

#### 4.6 Conclusions

Our study provided clinician's with a CDG involving findings from the manual spinal examination, palpation for segmental tenderness, and the extension-rotation test that is easy to implement in practice and enhances their decision-making ability when contemplating referral of their patients for facet joint interventions. Negative findings on the manual spinal examination and/or palpation for segmental tenderness may inform clinicians that facet joint blocks may not be an optimal management option for their patients with persistent neck pain. Future research in independent samples is needed to validate our CDG's prior to it's routine use in clinical practice.

## **Chapter 5: Derivation of a positive clinical decision guide of when NOT to refer patients for cervical facet joint blocks**

### 5.1 Background

Neck pain is a common clinical condition encountered by medical and allied health care professionals. Neck pain is typically episodic in nature, affecting up to 65% of individuals in their lifetime.<sup>1,61</sup> In the United States, self-reported spine pain and associated medical expenditures have increased in recent years, coinciding with a reduction in population health status.<sup>87</sup> This finding highlights the socioeconomic impact of spine pain in the general population. Neck pain can stem from various structures including the intervertebral disc, nerve root, facet, and ligaments.<sup>88</sup> Although specific etiology can be difficult to determine, studies using comparative, controlled facet joint blocks implicate the facet joint as a primary source of pain in 36% to 67% of those with persistent neck pain.<sup>6-8</sup> Patients undergoing facet joint interventions may be distinguished from patients with “typical mechanical neck pain” by their reports of persistent symptoms for at least three months duration, with higher levels of neck pain and disability, and a failure to respond to conservative rehabilitation and pharmacological interventions.<sup>5</sup>

The approach for diagnosis of facet joint mediated pain most recognized internationally are controlled block procedures using either two different local anaesthetics or placebo-controlled procedures.<sup>37 62</sup> It has been reported that the use of facet joint procedures in the United States increased by 624% between 1997 and 2006.<sup>13</sup> Facet joint blocks are invasive procedures, associated with significant costs and a small element of risk to the patient. There are lengthy wait-times for these procedures in many jurisdictions, where resources are limited. Patients who

ultimately respond negatively to diagnostic blocks magnify these wait-times. A clinical method to screen for patients most likely to benefit from diagnostic facet blocks would aid in reducing healthcare costs and wait-times. Psychological factors may play a role in the outcome of diagnostic facet joint blocks. There is some evidence suggesting that depression or high levels of psychopathology may influence the response to facet joint interventions, but the data is limited.<sup>58,101</sup> There is little evidence to suggest that any one factor related to the patients' history or clinical examination can predict the outcome of facet block procedures.<sup>15</sup> Thus, it has been suggested that the derivation of a clinical decision guide, incorporating findings from a cluster of clinical tests, may provide the clinician with a more accurate determination of those who may respond positively or negatively to diagnostic facet joint blocks.<sup>63</sup>

Although still controversial, findings from clinical tests such as range of motion, segmental palpation, the extension-rotation test, and manual spinal examination are used as guides to assist clinicians in making management decisions in the context of cervical facet joint mediated pain.<sup>14,51,63</sup> When examining the accuracy of clinical tests in diagnosing facet joint pain, one needs to consider tests that possess target specificity in identifying a particular spinal level as a potential pain generator. This is necessary since diagnostic facet joint block procedures target specific facet joints and is not applied generally to all facet joints in the neck for each patient.



## 5.2 Purpose

The purpose of this study was to derive a clinical decision guide (CDG) to identify patients that are not suitable for diagnostic facet joint blocks. In particular, our aim was to identify the predictive ability of psychological, behavioral, and physical factors related to a negative response to diagnostic facet joint blocks in order to inform clinicians when a facet joint intervention pathway may not be an appropriate management option for those with persistent neck pain.

## 5.3 Methods

### 5.3.1 *Study Design*

This study was a secondary analysis of a prospective cohort study conducted in a tertiary interventional pain management centre in Calgary, Alberta, Canada.

### 5.3.2 *Participants*

Consecutive patients with persistent neck pain, referred for diagnostic facet joint blocks were approached to participate. Subjects were included if aged between 18 and 65 years and reported neck pain intensity of  $\geq 3$  out of 10 on a Numeric Pain Rating Scale (NPRS) for at least the last three months.<sup>65</sup> This standard was set to ensure that a subject's pain intensity exceeded that of the reported measurement error of the NPRS.<sup>66</sup> Subjects were excluded if they presented with: cervical radiculopathy and/or upper motor neuron disease; neck pain related to systemic disease, infection, neoplasm, or fracture; a medically diagnosed psychological disorder; uncontrolled diabetes; uncontrolled clotting disorder; pregnancy; a workers compensation claim or ongoing litigation.

Consecutive sampling methods were applied and of the 177 individuals approached to participate in the study, 38 were excluded, 14 declined participation, and 125 consented to participate. Refer to Figure 1 (Appendix D) for a description of subject recruitment and participation. There were no clinically relevant differences in age, gender, neck pain intensity, and duration of neck pain between individuals who participated in the study and those who declined participation. Ethical approval for this study was obtained from the Conjoint Health Research Ethics Board at the University of Calgary (Ethics ID: 23783).

### *5.3.3 Procedures*

Once written consent was obtained, subjects completed a baseline demographic questionnaire followed by four standardized self-report measures. The Neck Disability Index (NDI) was completed to assess self-reported neck pain and disability.<sup>67</sup> Higher scores on the NDI reflect greater self-reported disability related to neck pain. The NDI has been shown to have a high degree of test-retest reliability (ICC: 0.89-0.94) and internal consistency (coefficient alpha values: 0.80-0.87).<sup>67</sup> The NDI has been reported to be a valid and responsive measure in those with mechanical neck disorders.<sup>66,89,90</sup> The Pain Catastrophizing Scale (PCS) was used to evaluate catastrophic thoughts in response to the patients neck pain.<sup>59</sup> From a clinical perspective, the PCS may be a useful tool in identifying patients that may be susceptible to adverse pain responses following invasive medical procedures. The PCS has demonstrated excellent internal consistency and moderate to substantial test-retest reliability.<sup>59,60</sup> The General Health Questionnaire – 28 (GHQ-28) is a reliable and valid measure of psychological distress in medical settings, and

it possesses four subscales.<sup>68,91</sup> The self-report version of the Leeds Assessment of Neuropathic Symptoms and Signs pain scale (S-LANSS) is a measure identifying pain of predominantly neuropathic origin.<sup>69</sup> The S-LANNS has demonstrated construct validity and excellent test-retest reliability.<sup>69,92</sup> Refer to Appendix A for a description of the self-report measures used in this study.

Subjects were assessed prior to their first diagnostic facet joint block. The subjects underwent a standardized clinical examination, which was performed by four experienced physiotherapists with 10-25 years of clinical experience. Select clinical tests commonly used in patients with neck pain were included in the standardized clinical examination. The physiotherapists were provided with a training manual outlining the standardized approach to the clinical examination, including operational definitions of the clinical tests. They underwent a one-hour training session to ensure a standardized approach. The physiotherapists were blinded to the level of the facet joint block to be performed for each subject. Data collection took place between October 2011 and December 2012.

The clinical examination was performed in the following sequence and included: the assessment of cervical range of motion (ROM), extension-rotation test, manual spinal examination, and palpation for segmental tenderness. Each testing session lasted approximately 15 minutes. Assessment of cervical range of motion, extension-rotation test, manual spinal examination, and palpation for segmental tenderness have demonstrated moderate to excellent intra-rater and inter-rater reliability in patients with axial neck pain referred for diagnostic facet joint blocks.<sup>93</sup>

#### 5.3.4 Outcome Measures

As per previous published research, the extension-rotation test, manual spinal examination, and palpation for segmental tenderness was performed to detect the presence or absence of cervical facet joint pain and dysfunction.<sup>14,51,71</sup> Refer to Appendix B for operational definitions of the clinical tests.

##### Criterion standard diagnostic facet joint block

Comparative medial branch blocks (MBB) were performed as the criterion standard for the diagnosis of facet joint pain.<sup>37</sup> Comparative blocks have been shown to minimize the risk of false positive responses reported by single block procedures.<sup>42</sup> This procedure involved the injection of 0.5 ml of either Bupivacaine 0.5% or Lidocaine 2% under fluoroscopic guidance onto the sensory nerves (medial branch of the dorsal ramus and/or the third occipital nerve) of the target facet joint. Contrast material (0.25 ml of Omnipaque 300) was injected at each spinal level to ensure target specificity of the facet blocks. All subjects underwent the criterion standard. The anaesthetics were delivered in a random order. The subjects and the interventional radiologist or physiatrist performing the injection were blinded to the anaesthetic used. The interventional radiologist, physiatrist, and radiology technician were blinded to any study related pre-injection outcomes measures or clinical examination findings.

All subjects underwent an initial MBB. A radiology technician recorded the subjects' neck pain intensity on an 11-point NPRS before and after the block. A positive response was defined as a  $\geq 80\%$  decrease in familiar neck pain intensity for at least the duration of the anaesthetic used ( $\geq 1$  hour for Lidocaine and 3 hours

for Bupivacaine).<sup>41</sup> Positive responders underwent a second MBB one week following their initial block. A subject was deemed to have facet joint pain if they were defined as a positive responder to both MBB's. If the subject reported < 80% relief in familiar neck pain intensity following the MBB, they were defined as a negative responder. In the negative responders, if the interventional radiologist or physiatrist performing the injection, or the referring physician felt that another facet joint is likely the putative source of neck pain, the subject may have undergone another facet joint block at a neighboring spinal level.

### 5.3.5 Data Analysis

The sample size was determined *a priori* based on the reported prevalence (36%-67%) of facet joint pain in the cervical spine.<sup>6-8</sup> With conservative use of this data, we estimated the prevalence of facet joint pain in our sample to be 40%. In deriving a CDG from multivariable regression analyses, it has been stated that at least 10 outcome events (diagnosis of facet joint pain) should occur for each predictor variable.<sup>22,94</sup> *A priori*, we determined that a CDG with more than five predictor variables may not be efficient for clinicians utilizing the guide in practice. From this, our study would require at least 50 positive outcome events (subjects diagnosed with facet joint pain). Given an anticipated prevalence of facet joint pain in our sample of 40% and the use of up to five predictor variables in the CDG, the number of subjects needed in our study was 125 ( $50/n = 0.40$ ). For this study, we adopted this sample size as it was used in our primary analysis for the derivation of a CDG that would inform clinicians on those who should respond positively to diagnostic facet joint blocks. Based on the premise that the number of negative

responders would be greater than the number of positive responders, we believed that this secondary analysis should be powered appropriately to meet the purpose of this study.

Descriptive statistics were used to summarize the subjects' baseline demographic data. Clinical tests, demographic, and psychological/behavioural measures reported in the literature to be associated with the outcome of diagnostic cervical facet joint blocks were analyzed for their association with the outcome of diagnostic facet blocks in our study via univariate and multivariate logistic regression. In consideration of the target specificity needed to determine the appropriate facet joints to undergo diagnostic injections, the manual spinal examination and palpation for segmental tenderness tests were selected for regression analysis and the derivation of the CDG. The extension-rotation test was also selected as it has been examined as a potential test to identify those with facet joint pain.<sup>51</sup> Model development also involved the findings from the NDI, GHQ-28, and PCS along with the clinical tests mentioned above, to account for the subject's baseline disability, psychological distress, and catastrophic processing of an anticipated or actual painful experience (diagnostic facet joint blocks).

Clinical tests with acceptable levels of intra-rater and inter-rater reliability (kappa statistic  $\geq 0.60$  or intraclass correlation coefficient  $\geq 0.80$ ) were considered in the development of the CDG.<sup>25</sup> For the regression analyses, the beta coefficients (and their standard errors) in the various models were evaluated for the effects of collinearity.<sup>95</sup> In addition, contingency tables were constructed for these variables to evaluate the extent of their agreement. If collinearity was detected then the data

from clinically relevant variables and the diagnostic outcome of the facet joint blocks was entered into contingency tables and the sensitivity, specificity, and likelihood ratios of the CDG's were calculated, along with their 95% confidence intervals. The objective was to develop a CDG that possessed the highest possible sensitivity and negative likelihood ratio in determining those without facet joint pain, while maximizing specificity. Interpretation of the magnitude of the likelihood ratios followed the guide reported by Guyatt and colleagues<sup>52</sup>, whereby likelihood ratios greater than 10 or less than 0.1 infer large and often conclusive shifts from pre-test to post-test probability; likelihood ratios from five to 10 and 0.1 to 0.2 infer a moderate shift in probability; likelihood ratios from two to five and 0.2 to 0.5 infer small, but sometimes important changes in probability; and likelihood ratios from one to two and 0.5 to 1 infer a change in probability that is rarely useful. The presence of effect modification by age, gender, catastrophization, and psychological distress was evaluated via logistic regression analyses. In addition, confounding by age, gender, baseline neck pain and disability, catastrophization, and psychological distress was controlled for in the logistic regression analyses.

For clinical relevance, exploratory analyses were performed to examine the association and predictive ability of the NDI, GHQ-28, and PCS utilizing cut-scores for each measure. The scores related to the 75<sup>th</sup> percentile for each measure were used as the cut-score as the authors believed that, clinically, the patients scoring in the upper 25% for each self-report measure may represent a subgroup of patients with the highest risk for a negative outcome from the diagnostic facet blocks.

Receiver operating characteristic (ROC) curves for potentially valuable prediction models were constructed, along with their respective areas under curve (AUC). All statistical analyses were performed using STATA 11 (StataCorp LP, College Station, Texas. USA) statistical software.

#### 5.4 Results

The baseline demographic data for the subjects, including their median scores (range) on the self-report measures, are presented in Table 5.1. The C5-6, C6-7, and C2-3 facet joints were the most frequent joints to undergo diagnostic facet joint blocks, with a prevalence of 36%, 33%, and 23% respectively. Of the 125 subjects, 73 had negative responses to comparative medial branch blocks, which amounts to a prevalence (pre-test probability) of those without facet joint pain of 58% in this sample. Of those that were negative responders, 14 were negative at C2-3; 13 were negative at C3-4; three were negative at C4-5; 24 were negative at C5-6; and 31 were negative at C6-7. Of these patients, 12 were negative at two levels (ie; C2-3 and C3-4).

In analyzing the results of the multivariate logistic regression analysis, collinearity was evident in the model with the variables representing the manual



**Table 5.1: Baseline Demographics**

	<b>Median (Range)</b> <b>N = 125</b>
Age (years)	49 (21 - 65)
Gender	84 females, 41 males
Baseline neck pain intensity (NPRS 0-10)	6 (3 - 9)
Subjects (%) with baseline neck pain intensity $\geq 5/10$	82
Onset of pain (%)	
Traumatic	47
Gradual	46
Sudden	7
Motor vehicle collision (%), (n=53)	
Frontal impact	15
Side impact	26
Rear impact	59
Smoker (%)	22
Employed (%)	
Currently employed	63
On leave due to neck problem	10
Unemployed	11
Retired	11
Student	5
Duration of neck pain (months)	18 (3 - 216)
Neck Disability Index (0-50)	20 (3 - 38)
Subjects (%) with Neck Disability Index scores $\geq 15/50$ (self-reported moderate-severe neck pain and disability)	74
Pain Catastrophizing Scale (0-52)	13 (0 - 44)
General Health Questionnaire - 28 (0-84)	22 (8 - 65)
s-LANNS (0-12)	8 (0 - 19)

spinal examination and palpation for segmental tenderness due to the high agreement between their findings (Table 5.2). For this reason, univariate logistic regression was utilized to describe the association between the manual spinal examination, palpation for segmental tenderness, and the extension-rotation test and the outcome of diagnostic facet joint blocks. In addition, contingency tables were generated and diagnostic accuracy statistics were calculated for these tests.

**Table 5.2: Contingency table of frequency of test findings for the manual spinal examination and palpation for segmental tenderness**

MSE	PST		Total
	Positive	Negative	
Positive	66*	3	69
Negative	3	53*	56
Total	69	56	125

Abbreviations: MSE = manual spinal examination, PST = palpation for segmental tenderness, \* = high agreement between the two clinical tests

From the univariate logistic regression analysis, the odds ratio associated with the NDI, GHQ-28, and the PCS was 0.98 (95% CI: 0.94 – 1.01), 0.98 (95% CI: 0.95 – 1.01), and 0.97 (95% CI: 0.94 – 1.01) respectively (Table 5.3). The odds ratio associated with the findings of the manual spinal examination, palpation for segmental tenderness, and the extension-rotation test was 29.71 (95% CI: 9.51 – 92.81), 43.28 (95% CI: 12.11 – 154.77), and 6.85 (95% CI: 2.91 – 16.13) respectively (Table 5.3).

For the manual spinal examination and palpation for segmental tenderness, negative findings occurred in 56 subjects (Table 5.4 and Table 5.5). For the extension-rotation test, negative findings occurred in 52 subjects (Table 5.6). Fifty-

**Table 5.3: Odds Ratios of the clinical variables in univariate logistic regression for predicting cervical facet joint pain**

<b>Clinical Variable</b>	<b>Coefficient</b>	<b>Standard Error</b>	<b>Odds Ratio (95% CI)</b>
Neck Disability Index	-0.08	0.03	0.98 (0.94 – 1.01)
General Health Questionnaire – 28	-0.02	0.02	0.98 (0.95 – 1.01)
Pain Catastrophizing Scale	-0.03	0.02	0.97 (0.94 – 1.01)
Extension-rotation test	1.92	0.44	6.85 (2.91 – 16.13)
Manual spinal examination	3.39	0.58	29.71 (9.51 – 92.81)
Palpation for segmental tenderness	3.77	0.65	43.28 (12.11 – 154.77)

Abbreviations: CI = confidence interval

**Table 5.4: Contingency table of frequency of test findings for the manual spinal examination and the outcome of diagnostic facet joint blocks**

MSE	Diagnostic facet joint block		Total
	Positive	Negative	
Positive	48	21	69
Negative	4	52	56
Total	52	73	125

Abbreviations: MSE = manual spinal examination

**Table 5.5: Contingency table comparing frequency of test findings for the palpation for segmental tenderness and the outcome of diagnostic facet joint blocks**

PST	Diagnostic facet joint block		Total
	Positive	Negative	
Positive	49	20	69
Negative	3	53	56
Total	52	73	125

Abbreviations: PST = palpation for segmental tenderness

**Table 5.6: Contingency table comparing frequency of test findings for the extension-rotation test and the outcome of diagnostic facet joint blocks**

ER	Diagnostic facet joint block		Total
	Positive	Negative	
Positive	43	30	73
Negative	9	43	52
Total	52	73	125

Abbreviations: ER = extension-rotation test

**Table 5.7: Contingency table comparing frequency of test findings for a combination of the MSE and PST and the outcome of diagnostic facet joint blocks**

MSE and PST	Diagnostic facet joint block		Total
	Positive	Negative	
Positive	48	18	66
Negative	4	55	59
Total	52	73	125

Abbreviations: MSE = manual spinal examination, PST = palpation for segmental tenderness

**Table 5.8: Contingency table comparing frequency of test findings for a combination of the MSE, PST, and ER test and the outcome of diagnostic facet joint blocks**

MSE/PST/ER	Diagnostic facet joint block		Total
	Positive	Negative	
Positive	41	12	53
Negative	11	61	72
Total	52	73	125

Abbreviations: MSE = manual spinal examination, PST = palpation for segmental tenderness, ER = extension-rotation test

nine subjects tested negative on both the manual spinal examination and palpation for segmental tenderness (Table 5.7), whereas, 72 subjects tested negative on all three clinical tests (Table 5.8).

If a subject tested negative on the manual spinal examination, the negative likelihood ratio was 0.11 (95% CI: 0.04 – 0.28) and the post-test probability of a diagnosis of facet joint pain decreased to 7% from the pre-test probability of 42%. Similarly, if a subject tested negative on palpation for segmental tenderness, the negative likelihood ratio was 0.08 (95% CI: 0.03 – 0.24) and the post-test probability of a diagnosis of facet joint pain decreased to 5% from 42%. If the subject tested negative on both clinical tests, the negative likelihood ratio was 0.10 (95% CI: 0.04 – 0.26) and the post-test probability of a diagnosis of facet joint pain decreased to 17% from 42%. Refer to Table 5.9 for the accuracy statistics associated with the CDG's.

In examining the association between multiple predictor variables and the outcome of the diagnostic facet joint blocks, we performed multivariate logistic regression analyses for two separate models (due to the presence of collinearity)

**Table 5.9: Accuracy statistics with 95% confidence intervals for the clinical decision guides**

Clinical Decision Guide	Sensitivity* (95% CI)	Specificity* (95% CI)	Negative Likelihood Ratio (95% CI)	Post-test Probability of a Diagnosis of Facet Joint Pain*†‡
CDG 1: MSE	92 (88 – 97)	71 (63 – 79)	0.11 (0.04 – 0.28)	7
CDG 2: 2. PST	94 (90 – 98)	73 (65 – 80)	0.08 (0.03 – 0.24)	5
CDG 3: ER	83 (76 – 89)	59 (50 – 68)	0.29 (0.16 – 0.54)	17
CDG 4: Subjects who tested negative on both MSE and PST	92 (88 – 97)	75 (68 – 83)	0.10 (0.04 – 0.26)	7

Abbreviations: CDG = clinical decision guide, MSE = manual spinal examination, PST = palpation for segmental tenderness, ER = extension-rotation test, CI = confidence interval

\* = proportions are stated as a percentage (%)

† = assumes a pre-test probability of 42%

‡ = post-test probability given a negative index test is equivalent to 1 – negative predictive value (for the index test or cluster of index tests)

with one containing the predictor variable palpation for segmental tenderness, and the other containing the predictor variable manual spinal examination. Multivariate logistic regression analysis, controlling for age, gender, catastrophization, baseline neck pain and disability, and psychological distress, examining the association between the findings of the palpation for segmental tenderness and the extension-rotation test revealed odds ratios of 40.99 (95% CI: 10.19 – 164.87) and 4.47 (95% CI: 1.34 – 14.98) respectively (Appendix E). There was no evidence of effect modification or confounding. For this model, ROC analysis resulted in an AUC of 0.90 and a sensitivity and specificity of 89% and 84% respectively (Appendix F). Utilizing the same analytical procedures for a model containing the findings of the manual spinal examination and the extension-rotation test the odds ratios reported were 28.57 (95% CI: 8.02 – 101.75) and 3.96 (95% CI: 1.23 – 12.79) respectively (Appendix G). Similarly, there was no evidence of effect modification or confounding. For this model, the AUC determined from ROC analysis was 0.89 and the sensitivity and specificity was 87% and 82% respectively (Appendix H).

For the exploratory analysis, the following cut scores for the self-report measures were used: PCS 24/52, NDI 24/50, and the GHQ-28 32/84. Using these cut-scores, the exploratory analysis involving univariate logistic regression demonstrated odds ratios of 0.46 (95% CI: 0.19 – 1.09), 0.63 (95% CI: 0.27 – 1.43), and 1.15 (0.52 – 2.55) for the GHQ-28, PCS, and NDI respectively. From contingency tables, the sensitivity and specificity of the GHQ-28 was 17% (95% CI: 11% - 24%) and 68% (95% CI: 60% - 77%) respectively. Similarly, the sensitivity and specificity for the PCS was 21% (95% CI: 14% - 28%) and 70% (95% CI: 62% - 78%)

respectively. Lastly, the NDI revealed a sensitivity of 29% (95% CI: 21% - 37%) and specificity of 74% (95% CI: 66% - 82%). The negative likelihood ratios associated with the GHQ-28, PCS, and NDI were 1.21 (95% CI: 1.05 – 1.38), 1.12 (95% CI: 0.97 – 1.31), and 0.96 (95% CI: 0.80 – 1.15) respectively.

## 5.5 Discussion

This study is the first study to derive a CDG, incorporating the findings from a cluster of clinical tests, informing clinicians when it may not be advantageous to direct patients toward cervical diagnostic facet joint blocks. Our results showed that the manual spinal examination and palpation for segmental tenderness exhibit high sensitivity (92% - 94%), in isolation or in combination (CDG 1, 2, or 4), supporting their use as potential screening tests prior to considering referring patients for facet joint blocks. As noted previously in Chapter 4, specificity (84%) of the CDG is maximized when a patient tests positive on the manual spinal examination, palpation for segmental tenderness, and the extension-rotation test. Coinciding with a test's sensitivity, the negative likelihood ratio is a useful statistic for quantifying a shift in probability favoring the condition of interest when the clinical test is negative.<sup>96</sup> In the context of our study, a clinical test, or cluster of clinical tests, that possess a low negative likelihood ratio would increase clinician confidence in making a decision of when not to refer a patient for a diagnostic facet joint block. Our data indicated that the lowest negative likelihood ratio (.08) was associated with the findings from the palpation for segmental tenderness test (CDG 2), suggesting that this test may be the most useful screening tool. The magnitude of this negative likelihood ratio provides clinicians with a large, and often conclusive,



shift in probability that the patient does not have facet joint pain, given a negative test finding.<sup>52,97</sup> In our study, knowledge of this negative likelihood ratio permitted a shift in post-test probability of the patient having facet joint pain to 5% from the pre-test probability of 42%. Similarly, the negative likelihood ratio associated with the manual spinal examination (CDG 1), and the combination of test findings from the manual spinal examination and the palpation for segmental tenderness test (CDG 4) was 0.11 and 0.10 respectively, generating a moderate change in probability shifting the post-test probability to 7%.

The manual spinal examination, palpation for segmental tenderness, and extension-rotation test were included in the derivation of the CDG's as there has been some literature indicating their use in evaluating those with potential cervical facet joint pain.<sup>14,15,51</sup> These tests aim to identify specific dysfunctional spinal motion segments. In clinical practice, diagnostic facet joint blocks target putative painful facet joints. The aim of these injections is not to target every facet joint in the cervical spine in every patient with persistent neck pain, as this would not be resourceful. Hence, clinical tests that possess target specificity are appropriate for consideration in our CDG aiming to identify those would respond negatively to facet joint blocks.

From our data, multivariate regression analyses determined that two models, one consisting of the manual spinal examination and extension-rotation test and the other involving palpation for segmental tenderness and extension-rotation test, (both models controlling for covariates related to baseline neck pain intensity, pain catastrophization, psychological distress, age, and gender) possessed good

sensitivity and specificity. In addition, the discriminative ability of both models, as reflected by the AUC from the ROC analysis, was satisfactory. However, the use of these multivariate models may not be practical in the clinical environment. Most clinicians may not have the statistical knowledge and/or technology to run such analyses to provide the odds ratio(s) of the patient having facet joint pain. The time necessary to input the variables into a regression program would also yield concern for the clinician. It has been said that clinical decision guides should be sensible, with the ultimate goal of changing clinician behaviour.<sup>102,103</sup> In our study, the manual spinal examination and palpation for segmental tenderness tests are relatively simple to use and the standardized testing protocol is efficient from a time perspective. In addition, these tests revealed high sensitivity and low negative likelihood ratios; therefore we emphasize the use of CDG's one, two, and four outlined in Table 5.9.

Of interest, our data did not identify a significant association between the self-report measures related to psychological distress, pain catastrophization, and baseline neck pain and disability and the outcome of the diagnostic facet joint blocks. Furthermore, from our exploratory analysis, the use of these measures with the pre-determined cut-points used in our study, does not support their use as screening tools to determine which patients may have a negative response to the facet blocks. Using similar diagnostic block procedures as reported in our study, Manchikanti et al.<sup>101</sup> studied the influence of psychopathology (depression, generalized anxiety disorder, and somatization, alone or in combination), diagnosed as per DSM-IV-TR criteria, on the prevalence of cervical, thoracic, and lumbar facet

joint pain and the false positive rate of single facet joint blocks. The authors found a higher prevalence of cervical facet joint pain along with a lower false positive rate with single blocks in those with major depression. By using the criterion standard DSM-IV-TR criteria and physician interview to formulate the diagnosis of major depression, it provides a rationale as to why their study findings differed from ours. In our study, we administered the GHQ-28 as a measure of psychological distress, and although it has a subscale related to severe depression, it is not as sensitive as the DSM-IV-TR criteria in classifying patients with major depression. In a related study, Wasan et al.<sup>58</sup> found that higher levels of psychopathology (anxiety and depression) was associated with a worsening of neck pain intensity in patients one month following a single medial branch block with corticosteroids. Their findings should be interpreted with caution, as the authors did not use comparative, controlled facet joint blocks. They incorporated single facet blocks, known to possess a high false positive rate ranging from 27% to 63%, possibly leading to a misclassification of those determined to have facet joint pain.<sup>42,43</sup> From this one cannot conclude that higher levels of psychopathology are associated with poorer outcomes from facet joint blocks.

Our study is the first to evaluate the accuracy of the extension-rotation test in the diagnosis of cervical facet joint pain. Our data indicated that if a patient tests negative on this test (sensitivity = 83% and negative likelihood ratio = 0.29), this test result alone might not substantiate a clinician's diagnosis that the patient does not have facet joint mediated pain.

Siegenthaler et al.<sup>98</sup> assessed mechanical pain sensitivity (pressure pain thresholds measured in kPa), via an electronic pressure algometer, in 33 patients with unilateral neck pain. They hypothesized that the pain threshold would be lower over the painful facet joint compared with the contralateral side, attempting to validate segmental tenderness as clinical test in identifying facet joint pain. Using a cut-point of 1 kPa (mean difference between the painful side and non-painful side) to categorize the pressure algometry as positive or negative and comparing these findings with the outcome of comparative, controlled medial branch blocks, the authors reported a sensitivity of 67% and specificity of 16%. As the authors did not report the negative likelihood ratio, we calculated it to be 2.06. Using a much higher cut-point of 30 kPa, the tests sensitivity decreased to 13%, specificity increased to 95% and the negative likelihood ratio decreased to 0.92. The use of dichotomous cut-points can be problematic in prediction research, as it has been shown to diminish the discriminative ability of the model or test.<sup>99</sup> In our study, the assessors recorded a positive test for segmental tenderness when the subject reported “familiar pain” along with a cut-off of neck pain intensity of  $\geq 3/10$  on a NPRS. The addition of “familiar pain” provocation in the operational definition of the test may aide in distinguishing whether a particular spinal motion segment is the specific source of pain. Furthermore, the use of the clinician’s hands directly palpating taut bands of segmental muscle (compared to an electronic device) may have improved the tests accuracy which is supported by the neuroanatomical relationship between the deep cervical segmental muscles and the facet joint capsule.<sup>39,72</sup> When interpreting the precision of the findings of Siegenthaler et al.<sup>98</sup>, one needs to be

cautious as the study was not powered to determine diagnostic accuracy statistics. In consideration of the small sample size, the confidence intervals associated with the accuracy estimates (which the authors did not report) would be wide questioning the precision of estimated sensitivity, specificity, and negative likelihood ratio. In addition, the outcome of pressure pain threshold testing in patients with neck pain is known to be highly variable.<sup>100</sup> Using a mean difference in kPa to formulate the cut-point for a positive test is problematic when dealing with highly variable outcomes, presenting another reason to be cautious when interpreting the reliability of their accuracy statistics.

A landmark study by Jull et al.<sup>14</sup> examined the ability of an experienced physiotherapist to identify symptomatic facet joints in 20 patients with neck pain undergoing single diagnostic facet joint blocks, which was the criterion standard at the time of their study. Incorporating a manual spinal examination, the physiotherapist correctly identified those with and without facet joint pain, along with the specific segmental level of pain in all 20 patients. Remarkably, the manual spinal examination used in their study was 100% sensitive and 100% specific in deciphering those with and without facet joint pain. In our study, we incorporated a similar manual spinal examination procedure, but we compared our index test findings against those of the currently accepted criterion standard, comparative, controlled medial branch blocks, for the diagnosis of facet joint pain.<sup>37</sup> In the case of single diagnostic blocks, we now know that they possess a high false positive rate.<sup>42,43</sup> This finding undermines the magnitude of the sensitivity and specificity

noted in the study by Jull et al.<sup>14</sup>, and may explain some of the discrepancies between our study findings.

Using comparative, controlled medial branch blocks as the criterion standard, King et al.<sup>15</sup> studied the validity of the manual spinal examination in assessing patients with neck pain. The manual spinal examination mimicked the procedure used by Jull et al.<sup>14</sup> The authors reported a sensitivity of 89% (95% CI: 82% - 96%), specificity of 47% (95% CI: 37% - 57%), and a positive likelihood ratio of 1.7 (95% CI: 1.2 - 2.5) questioning the criterion validity of the manual spinal examination. As the authors did not provide the negative likelihood ratio, using their values for sensitivity and specificity, we calculated it to be 0.23 (95% CI: 0.13 - 0.42). Although the estimated test sensitivity was similar to that of our study, King et al.<sup>15</sup> noted that only patients with “clinically positive” joints (secondary to the results of the manual spinal examination) underwent the criterion standard facet joint blocks. If the criterion standard is provided solely to patients that test positive on the index test, measurement bias (verification bias) may affect the accuracy statistics of the index test under study. Verification bias may explain the differences, and caution assumptions around the similarities, in our study findings compared with that of King et al.<sup>15,52</sup> This form of bias may have lead to an underestimate of the specificity and overestimate of the sensitivity of the manual spinal examination noted in the study by King et al.<sup>15,53</sup> Nonetheless, our data suggests that the manual spinal examination may be considered as a possible screening tool for ‘ruling out’ those with neck pain of facet joint origin.

### *5.5.1 Study Limitations*

There are potential limitations in this study. Our assessors were experienced physiotherapists, thus the findings may not necessarily be generalizable to all clinicians. However, the comprehensive, time efficient, and standardized approach to the clinical tests utilized in this study fosters ease of use in clinical practice.<sup>93</sup>

The subjects underwent diagnostic facet joint blocks to the probable painful joints based on their pain pattern and response to the previous facet joint block. Since they did not receive facet blocks at every spinal level in the neck, some might suggest that this compromises the validity of our diagnostic accuracy statistics. Although this raises a methodological concern, it is neither realistic nor efficient, from a clinical perspective to put every patient through this number of injections. By having every subject undergo the criterion standard, we believe that we minimized bias related to diagnostic work-up.<sup>53</sup>

The CDG's in our study were derived from the combination of tests that optimized sensitivity and the negative likelihood ratio calculated from contingency tables. In recent years, the use of multivariate regression modeling to formulate a CDG has provided a method of predicting an outcome from multiple risk factors while simultaneously evaluating effect modification and controlling for confounding. In our study, we reported the outcome of multivariate regression analyses fitting two separate models, one using the manual spinal examination and extension-rotation test as primary exposure variables and the other using the palpation for segmental tenderness and extension-rotation tests as the primary exposure variables. Although it may have been interesting to model the manual spinal

examination and palpation for segmental tenderness test simultaneously, our decision to perform the multivariate analyses with the tests in two separate models was based on the collinearity observed between the two tests when assessed in the same model. The existence of collinearity may inflate the variances related to the estimated coefficients in a model resulting in a lack of statistical significance of one or more independent variables, although the overall model may be significant.<sup>95</sup> Although one can combine highly correlated variables into a single variable for analysis in a multivariate model, we believe that the constructs underlying the manual spinal examination and palpation for segmental tenderness is distinctive for each test, as the former test has involves an assessment of mobility. In our view, this renders the information obtained from each test clinically relevant and unique in it owns right. In this context, reporting of univariate regression analyses was informative, while avoiding erroneous reporting of its multivariate counterpart.

This study may have omitted potential predictor variables that could have been incorporated in the derivation of the CDG's. We chose to include clinical tests that have either been studied previously for their diagnostic accuracy in those with cervical facet joint pain or have been widely used by clinician's based on the assumption of the ability of the test to provoke facet joint pain. We avoided including multiple predictor variables that did not have a biologically plausible association with facet joint pain, to minimize the risk of analyzing multiple variables without enough power to detect important predictors.

Additionally, the measure used to reflect psychological distress, although it possesses a subscale to evaluate severe depression, it may not have been sensitive



enough to classify those with major depression. Subsequently, the use of a different method to measure the presence of depression may provide a more accurate determination of this variable as a predictor of the outcome of cervical facet joint blocks.

## 5.6 Conclusions

Our study provides clinicians with CDG's utilizing findings from clinical tests that are relatively easy to implement in practice and may enhance their decision-making ability when contemplating referral of their patients for facet joint interventions. Negative findings on the manual spinal examination and/or palpation for segmental tenderness can inform clinicians that facet joint blocks may not be an optimal management option for their patients with persistent neck pain. Future research in independent samples is needed to validate our CDG's prior to its routine use in clinical practice.

## Chapter 6: Conclusions

Neck pain research has consumed the published scientific literature over the past two decades. The heterogeneous nature of neck pain fosters a challenge for clinicians when formulating a diagnosis and directing intervention pathways. This thesis provided insight into current gaps in clinical research, by examining the reliability and diagnostic accuracy of standardized clinical tests in the diagnosis of cervical facet joint pain. As this research is at the derivation stage of a clinical decision guide, it provides a platform for future validation research in independent samples of patients with persistent neck pain.

Chapter two provided a narrative review of the literature, outlining the current state of the use of facet joint interventions along with the challenges facing clinicians in terms of the diagnostic accuracy of individual clinical tests in the diagnosis of cervical facet joint pain. Gaps in current research were identified, highlighting the need for research examining the diagnostic utility of a cluster of clinical tests in the diagnosis of cervical facet joint pain. The use of results from a cluster of test findings mirrors clinical practice and supports clinical reasoning processes.<sup>26,54</sup>

The use of facet joint interventions has increased exponentially over the last two decades. This finding is reflected in lengthy wait-lists and economic implications associated with facet joint procedures. Clinicians are faced with the challenge and responsibility of making appropriate referrals for these procedures. In the absence of the criterion standard for the diagnosis of facet joint pain in the practices of family physicians, physiotherapists, and other primary care providers,

these health care professionals need to rely on reliable and valid clinical tests to decipher patients that should undergo facet joint procedures. In Chapter 3, the intra-rater and inter-rater reliability of select clinical tests in 56 patients referred for diagnostic facet joint blocks was reported. The standardized testing protocol exhibited moderate to excellent reliability. Based on the methodological standards suggested for the inclusion of clinical tests in the development of CDG's, it was determined that measurement of cervical spine range of motion, the extension-rotation test, the manual spinal examination, and palpation for segmental tenderness could be included in the derivation of a CDG for the diagnosis of cervical facet joint pain.<sup>25</sup>

Subsequently, CDG's were developed in Chapter 4 informing clinicians on when to refer patients with persistent neck pain for diagnostic facet joint blocks. From this, a cluster of test involving the manual spinal examination, palpation for segmental tenderness, and the extension-rotation test possessed a specificity of 84% (95% CI: 77% - 90%) and a positive likelihood ratio of 4.94 (2.80 - 8.20). In our study, the pre-test probability of having facet joint pain was 42%. Incorporating the positive likelihood ratio of 4.94, if a patient tested positive on all three clinical tests this would create a shift in post-test probability of having facet joint pain to 78%.

Interestingly, the utility of the clinical tests outlined above may be highlighted as potential screening tools enabling clinicians to steer patients with persistent neck pain away from facet joint interventions. At the outset of this research, we did not strongly consider the manual spinal examination and palpation

for segmental tenderness test as potential screening procedures. Therefore, there was some modest redundancy in reporting noted between chapters four and five. Nonetheless, in keeping with the purpose of chapter five, we described a negative clinical decision guide identifying patients that may not be appropriate candidates for facet joint blocks. Interestingly, the palpation for segmental tenderness test and the manual spinal examination demonstrated a sensitivity of 94% (95% CI: 90% – 98%) and 92% (95% CI: 88% – 97%) respectively. In addition, the negative likelihood ratios associated with the palpation for segmental tenderness test and manual spinal examination was 0.08 (0.03 – 0.24) and 0.11 (0.04 – 0.28) respectively. From this, if a patient tests negative on the palpation for segmental tenderness test a shift in probability of the patient having facet joint pain from 42% to 5% would occur. Similar findings were found for the manual spinal examination.

In addition, multivariate regression analyses determined that two models, one consisting of the manual spinal examination and extension-rotation test and the other involving palpation for segmental tenderness and extension-rotation test, (both models controlling for covariates related to baseline neck pain intensity, pain catastrophization, psychological distress, age, and gender) possessed good sensitivity and specificity. In addition, the discriminative ability of both models, as reflected by the AUC from the ROC analysis, was satisfactory. However, the use of these multivariate models may not be practical in the clinical environment. The time necessary to input the variables into a regression program may be of concern for the clinician, leading them not to use this model in practice. As a consequence, we highlighted the use of the manual spinal examination and palpation for segmental

tenderness test as possible screening tools due to their advantage of clinical practicality. Exploratory analysis of the predictive ability of self-report measures (using pre-determined cut-points) related to psychological distress, pain catastrophization, and baseline neck pain and disability did not support the use of these measures as screening tools for cervical facet joint pain.

Our findings provide a foundation for future research to validate the CDG's in larger independent samples of individuals with persistent neck pain. As there is a possibility that a different set of predictor variables may transpire in different group of patients, validation of the CDG's is essential prior to their routine use in clinical practice.<sup>22</sup> Although our research findings are intriguing, consideration of other important predictors or measures is needed in future research. An argument can be raised that the application of clinical tests such as the manual spinal examination and palpation for segmental tenderness requires a clinician with specialized skills. The evidence to date does not refute that notion, thus validation research may involve practitioners with varying levels of experience.

The optimal CDG is one that alters clinician behaviour such that it is used effectively in clinical practice and results in improved patient satisfaction and outcomes.<sup>102,103</sup> Undoubtedly, this process is a large undertaking and may involve the solicitation of clinician input, further research evaluating the utility of the CDG's against clinician judgement, and implementation of effective knowledge translation strategies. Health economic research evaluating the cost-effectiveness of the use of the CDG's in practice compared to that of standard care is needed to demonstrate its impact and solidify its use amongst clinicians. All of these considerations will inform

future research evaluating the effectiveness of the CDG's with the ultimate aim of assisting clinicians in optimizing the quality of care for patients with persistent neck pain.

## Bibliography

1. Cote P, Cassidy JD, Carroll L. The Saskatchewan Health and Back Pain Survey. The prevalence of neck pain and related disability in Saskatchewan adults. *Spine*. Aug 1 1998;23(15):1689-1698.
2. Hogg-Johnson S, van der Velde G, Carroll LJ, et al. The burden and determinants of neck pain in the general population: results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. *Spine (Phila Pa 1976)*. Feb 15 2008;33(4 Suppl):S39-51.
3. Croft PR, Lewis M, Papageorgiou AC, et al. Risk factors for neck pain: a longitudinal study in the general population. *Pain*. Sep 2001;93(3):317-325.
4. Davis MA, Onega T, Weeks WB, Lurie JD. Where the United States spends its spine dollars: expenditures on different ambulatory services for the management of back and neck conditions. *Spine*. Sep 1 2012;37(19):1693-1701.
5. Manchikanti L, Helm S, Singh V, et al. An algorithmic approach for clinical management of chronic spinal pain. *Pain Physician*. Jul-Aug 2009;12(4):E225-264.
6. Lord SM, Barnsley L, Wallis BJ, Bogduk N. Chronic cervical zygapophysial joint pain after whiplash. A placebo-controlled prevalence study. *Spine*. Aug 1 1996;21(15):1737-1744; discussion 1744-1735.
7. Manchikanti L, Singh V, Pampati V, Damron KS, Beyer CD, Barnhill RC. Is there correlation of facet joint pain in lumbar and cervical spine? An evaluation of prevalence in combined chronic low back and neck pain. *Pain Physician*. Oct 2002;5(4):365-371.

- 8.** Manchikanti L, Manchikanti K, Pampati V, Brandon D, Giordano J. The prevalence of facet joint-related chronic neck pain in postsurgical and nonpostsurgical patients: A comparative evaluation. *Pain Practice*. 2008;8(1):5-10.
- 9.** Lord SM, Barnsley L, Wallis BJ, McDonald GJ, Bogduk N. Percutaneous radiofrequency neurotomy for chronic cervical zygapophyseal-joint pain. *N Engl J Med*. Dec 5 1996;335(23):1721-1726.
- 10.** McDonald GJ, Lord SM, Bogduk N. Long-term follow-up of patients treated with cervical radiofrequency neurotomy for chronic neck pain. *Neurosurgery*. Jul 1999;45(1):61-67; discussion 67-68.
- 11.** Barnsley L. Percutaneous radiofrequency neurotomy for chronic neck pain: outcomes in a series of consecutive patients. *Pain Med*. Jul-Aug 2005;6(4):282-286.
- 12.** MacVicar J, Borowczyk JM, MacVicar AM, Loughnan BM, Bogduk N. Cervical medial branch radiofrequency neurotomy in New Zealand. *Pain medicine*. May 2012;13(5):647-654.
- 13.** Manchikanti L, Singh V, Pampati V, Smith HS, Hirsch JA. Analysis of growth of interventional techniques in managing chronic pain in the Medicare population: a 10-year evaluation from 1997 to 2006. *Pain Physician*. Jan-Feb 2009;12(1):9-34.
- 14.** Jull G, Bogduk N, Marsland A. The accuracy of manual diagnosis for cervical zygapophysial joint pain syndromes. *The Medical Journal of Australia*. Mar 7 1988;148(5):233-236.
- 15.** King W, Lau P, Lees R, Bogduk N. The validity of manual examination in assessing patients with neck pain. *The Spine Journal*. 2007;7:22-26.



16. Smedmark V, Wallin M, Arvidsson I. Inter-examiner reliability in assessing passive intervertebral motion of the cervical spine. *Manual Therapy*. May 2000;5(2):97-101.
17. Piva SR, Erhard RE, Childs JD, Browder DA. Inter-tester reliability of passive intervertebral and active movements of the cervical spine. *Man Ther*. Nov 2006;11(4):321-330.
18. Pool JJ, Hoving JL, de Vet HC, van Mameren H, Bouter LM. The interexaminer reproducibility of physical examination of the cervical spine. *Journal of Manipulative and Physiological Therapeutics*. Feb 2004;27(2):84-90.
19. Cleland JA, Childs JD, Fritz JM, Whitman JM. Interrater reliability of the history and physical examination in patients with mechanical neck pain. *Archives of Physical Medicine and Rehabilitation*. Oct 2006;87(10):1388-1395.
20. Hall T, Briffa K, Hopper D, Robinson K. Reliability of manual examination and frequency of symptomatic cervical motion segment dysfunction in cervicogenic headache. *Manual Therapy*. Dec 2010;15(6):542-546.
21. Jull G, Zito G, Trott P, Potter H, Shirley D, Richardson C. Inter-examiner reliability to detect painful upper cervical joint dysfunction. *Aust J Physiother*. 1997;43(2):125-129.
22. McGinn T, Guyatt G, Wyer P, Naylor C, Stiell I, Richardson W. Users' guides to the medical literature XXII: How to use articles about clinical decision rules. *Journal of the American Medical Association*. 2000;284(1):79-84.

- 23.**Hebert JJ, Fritz JM. Clinical decision rules, spinal pain classification and prediction of treatment outcome: A discussion of recent reports in the rehabilitation literature. *Chiropr Man Therap.* 2012;20(1):19.
- 24.**Childs J, Cleland J. Development and application of clinical prediction rules to improve decision making in physical therapist practice. *Phys Ther.* 2006;86(1):122-131.
- 25.**Stiell IG, Wells GA. Methodologic standards for the development of clinical decision rules in emergency medicine. *Annals of Emergency Medicine.* Apr 1999;33(4):437-447.
- 26.**Cook C, Hegedus E. Diagnostic utility of clinical tests for spinal dysfunction. *Manual Therapy.* Feb 2011;16(1):21-25.
- 27.**Deyo RA, Weinstein JN. Low back pain. *New England Journal of Medicine.* Feb 1 2001;344(5):363-370.
- 28.**Lord SM, Bogduk N. Radiofrequency procedures in chronic pain. *Best Pract Res Clin Anaesthesiol.* Dec 2002;16(4):597-617.
- 29.**Manchikanti L, Boswell MV, Datta S, et al. Comprehensive review of therapeutic interventions in managing chronic spinal pain. *Pain Physician.* Jul-Aug 2009;12(4):E123-198.
- 30.**McLain R. Mechanoreceptors endings in human cervical facet joints. *Spine.* 1994;19(5):495-501.
- 31.**Kallakuri S, Singh A, Chen C, Cavanaugh J. Demonstration of substance P, Calcitonin gene-related peptide, and protein gene product 9.5 containing nerve fibers in human cervical facet joint capsules. *Spine.* 2004;29(11):1182-1186.

- 32.**Brain S. Sensory neuropeptides: Their role in inflammation and wound healing. *Immunopharmacology*. 1997;37:133-152.
- 33.**Ackermann PW, Li J, Lundeborg T, Kreicbergs A. Neuronal plasticity in relation to nociception and healing of rat achilles tendon. *Journal of Orthopaedic Research*. May 2003;21(3):432-441.
- 34.**Chen C, Lu Y, Kallakuri S, Patwardhan A, Cavanaugh J. Distribution of A-delta and C-fiber receptors in the cervical facet joint capsule and their response to stretch. *Journal of Bone and Joint Surgery*. 2006;88 - A(8):1807-1816.
- 35.**Dwyer A, Aprill C, Bogduk N. Cervical zygapophyseal joint pain patterns. I: A study in normal volunteers. *Spine*. Jun 1990;15(6):453-457.
- 36.**Aprill C, Dwyer A, Bogduk N. Cervical zygapophyseal joint pain patterns. II: A clinical evaluation. *Spine*. Jun 1990;15(6):458-461.
- 37.**Bogduk N. International Spinal Injection Society Guidelines for the performance of spinal injection procedures: Part 1: Zygapophysial joint blocks. *Clinical Journal of Pain*. 1997;13(4):292-297.
- 38.**Merskey H, Bogduk N. Classification of chronic pain. In: Merskey H, Bogduk N, eds. *Descriptions of chronic pain syndromes and definition of pain terms*. 2 ed. Seattle: IASP Press; 1994:180-181.
- 39.**Bogduk N. The clinical anatomy of the cervical dorsal rami. *Spine*. Jul-Aug 1982;7(4):319-330.
- 40.**Barnsley L, Bogduk N. Medial branch blocks are specific for the diagnosis of cervical zygapophyseal joint pain. *Regional Anesthesia and Pain Medicine*. Nov-Dec 1993;18(6):343-350.

41. Boswell MV, Singh V, Staats PS, Hirsch JA. Accuracy of precision diagnostic blocks in the diagnosis of chronic spinal pain of facet or zygapophysial joint origin. *Pain Physician*. Oct 2003;6(4):449-456.
42. Barnsley L, Lord S, Wallis B, Bogduk N. False-positive rates of cervical zygapophyseal joint blocks. *Clinical Journal of Pain*. 1993;9:124-130.
43. Manchikanti L, Boswell MV, Singh V, Pampati V, Damron KS, Beyer CD. Prevalence of facet joint pain in chronic spinal pain of cervical, thoracic, and lumbar regions. *BMC Musculoskelet Disord*. May 28 2004;5:15.
44. Falco FJ, Erhart S, Wargo BW, et al. Systematic review of diagnostic utility and therapeutic effectiveness of cervical facet joint interventions. *Pain Physician*. Mar-Apr 2009;12(2):323-344.
45. Niemisto L, Kalso E, Malmivaara A, Seitsalo S, Hurri H. Radiofrequency denervation for neck and back pain: a systematic review within the framework of the cochrane collaboration back review group. *Spine (Phila Pa 1976)*. Aug 15 2003;28(16):1877-1888.
46. Curatolo M, Reiz S. Re: Niemisto L, Kalso E, Malmivaara A, et al. Radiofrequency denervation for neck and back pain: a systematic review within the framework of the Cochrane Collaboration Back Review Group. *Spine* 2003;28:1877-88. *Spine (Phila Pa 1976)*. Jan 15 2005;30(2):263-264; author reply 264-265.
47. Husted DS, Orton D, Schofferman J, Kine G. Effectiveness of repeated radiofrequency neurotomy for cervical facet joint pain. *Journal of spinal disorders & techniques*. Aug 2008;21(6):406-408.

48. Rambaransingh B, Stanford G, Burnham R. The effect of repeated zygapophysial joint radiofrequency neurotomy on pain, disability, and improvement duration. *Pain medicine*. Sep 2010;11(9):1343-1347.
49. Cleland J, Childs J, Fritz J, Whitman J. Interrater reliability of the history and physical examination in patients with mechanical neck pain. *Archives of Physical Medicine and Rehabilitation*. 2006;87(10):1388-1395.
50. Fjellner A, Bexander C, Faleij R, Strender LE. Interexaminer reliability in physical examination of the cervical spine. *Journal of Manipulative and Physiological Therapeutics*. Oct 1999;22(8):511-516.
51. Cohen S, Bajwa Z, Kraemer J, et al. Factors predicting success and failure for cervical facet radiofrequency denervation: A multi-center analysis. *Regional Anesthesia and Pain Medicine*. 2007;32(6):495-503.
52. Guyatt G, Rennie D, Meade M, Cook D. *Users' Guides to the Medical Literature: A manual for evidence-based clinical practice*. 2 ed. Toronto: McGraw-Hill Companies Inc.; 2008.
53. Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *Journal of the American Medical Association*. Sep 15 1999;282(11):1061-1066.
54. Haldeman S. Is it time to discard the term "diagnosis" when examining a person with uncomplicated axial neck pain? *The Spine Journal*. 2011;11:177-179.
55. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. *Journal of the American Medical Association*. Feb 12 1997;277(6):488-494.

56. Anis AH, Stiell IG, Stewart DG, Laupacis A. Cost-effectiveness analysis of the Ottawa Ankle Rules. *Annals of Emergency Medicine*. Oct 1995;26(4):422-428.
57. Stiell IG, Clement CM, Grimshaw J, et al. Implementation of the Canadian C-Spine Rule: prospective 12 centre cluster randomised trial. *BMJ*. 2009;339:b4146.
58. Wasan A, Jamison R, Pham L, Tipirneni N, Nedeljkovic S, Katz J. Psychopathology predicts the outcome of medial branch blocks with corticosteroid for chronic axial low back or cervical pain: A prospective cohort study. *BMC Musculoskeletal Disorders*. 2009;10(22).
59. Sullivan M, Bishop S, Pivik J. The pain catastrophizing scale: Development and validation. *Psychological Assessment*. 1995;7:524-532.
60. Rhudy J, France C, Bartley E, Williams A, McCabe K, Russell J. Does pain catastrophizing moderate the relationship between spinal nociceptive processes and pain sensitivity? *Journal of Pain*. 2009;10(8):860-869.
61. Chiu TT, Leung AS. Neck pain in Hong Kong: a telephone survey on prevalence, consequences, and risk groups. *Spine*. Jul 15 2006;31(16):E540-544.
62. Lord SM, Barnsley L, Bogduk N. The utility of comparative local anesthetic blocks versus placebo-controlled blocks for the diagnosis of cervical zygapophysial joint pain. *Clin J Pain*. Sep 1995;11(3):208-213.
63. Schneider GM, Jull G, Thomas K, Salo P. Screening of patients suitable for diagnostic cervical facet joint blocks--a role for physiotherapists. *Manual Therapy*. Apr 2012;17(2):180-183.
64. Shrout PE. Measurement reliability and agreement in psychiatry. *Stat Methods Med Res*. Sep 1998;7(3):301-317.

65. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. *J Clin Nurs*. Aug 2005;14(7):798-804.
66. Cleland JA, Childs JD, Whitman JM. Psychometric properties of the Neck Disability Index and Numeric Pain Rating Scale in patients with mechanical neck pain. *Arch Phys Med Rehabil*. Jan 2008;89(1):69-74.
67. Vernon H, Mior S. The neck disability index: a study of reliability and validity. *J Manipulative Physiol Ther*. 1991;14:409-415.
68. Goldberg DP, Hillier VF. A scaled version of the General Health Questionnaire. *Psychol Med*. Feb 1979;9(1):139-145.
69. Bennett MI, Smith BH, Torrance N, Potter J. The S-LANSS score for identifying pain of predominantly neuropathic origin: validation for use in clinical and postal research. *J Pain*. Mar 2005;6(3):149-158.
70. van Trijffel E, Anderegg Q, Bossuyt PM, Lucas C. Inter-examiner reliability of passive assessment of intervertebral motion in the cervical and lumbar spine: a systematic review. *Manual Therapy*. Nov 2005;10(4):256-269.
71. Uthaikhup S, Sterling M, Jull G. Cervical musculoskeletal impairment is common in elders with headache. *Man Ther*. Dec 2009;14(6):636-641.
72. Winkelstein BA, McLendon RE, Barbir A, Myers BS. An anatomical investigation of the human cervical facet capsule, quantifying muscle insertion area. *J Anat*. Apr 2001;198(Pt 4):455-461.
73. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther*. Mar 2005;85(3):257-268.

74. Donner A, Eliasziw M. A goodness-of-fit approach to inference procedures for the kappa statistic: confidence interval construction, significance-testing and sample size estimation. *Stat Med.* Aug 1992;11(11):1511-1519.
75. Cohen J. *Statistical power analysis for the behavioural sciences*. 1st ed. London: Academic Press; 1977.
76. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* Mar 1977;33(1):159-174.
77. Portney L, Watkins M. *Foundations of clinical research: Applications to practice*. 3rd ed. Upper Saddle River, New Jersey: Prentice Hall Health; 2009.
78. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* Feb 8 1986;1(8476):307-310.
79. Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *BMJ.* Jun 6 1992;304(6840):1491-1494.
80. Wainner R, Fritz J, Irrgang J, Boninger M, Delitto A, Allison S. Reliability and diagnostic accuracy of the clinical examination and patient self-report measures for cervical radiculopathy. *Spine.* 2003;28(1):52-62.
81. Strender LE, Lundin M, Nell K. Interexaminer reliability in physical examination of the neck. *Journal of Manipulative and Physiological Therapeutics.* Oct 1997;20(8):516-520.
82. Feinstein A, Cicchetti D. High agreement but low Kappa: I. The problems of two paradoxes. *J Clin Epidemiol.* 1990;43(6):543-549.



- 83.**van Suijlekom H, de Vet HC, van den Berg S, Weber W. Interobserver reliability in physical examination of the cervical spine in patients with headache. *Headache*. 2000;40:581-586.
- 84.**Viikari-Juntura E. Interexaminer reliability of observations in physical examinations of the neck. *Phys Ther*. Oct 1987;67(10):1526-1532.
- 85.**Deboer KF, Harmon R, Jr., Tuttle CD, Wallace H. Reliability study of detection of somatic dysfunctions in the cervical spine. *Journal of Manipulative and Physiological Therapeutics*. Mar 1985;8(1):9-16.
- 86.**Hubka MJ, Phelan SP. Interexaminer reliability of palpation for cervical spine tenderness. *Journal of Manipulative and Physiological Therapeutics*. Nov-Dec 1994;17(9):591-595.
- 87.**Martin BI, Deyo RA, Mirza SK, et al. Expenditures and health status among adults with back and neck problems. *JAMA : the journal of the American Medical Association*. Feb 13 2008;299(6):656-664.
- 88.**Alexander EP. History, physical examination, and differential diagnosis of neck pain. *Physical medicine and rehabilitation clinics of North America*. Aug 2011;22(3):383-393, vii.
- 89.**Young IA, Cleland JA, Michener LA, Brown C. Reliability, construct validity, and responsiveness of the neck disability index, patient-specific functional scale, and numeric pain rating scale in patients with cervical radiculopathy. *American journal of physical medicine & rehabilitation / Association of Academic Physiatrists*. Oct 2010;89(10):831-839.

- 90.**Gustavsson C, Bergstrom J, Denison E, von Koch L. Predictive factors for disability outcome at twenty weeks and two years following a pain self-management group intervention in patients with persistent neck pain in primary health care. *Journal of rehabilitation medicine : official journal of the UEMS European Board of Physical and Rehabilitation Medicine*. Jan 25 2013;45(2):170-176.
- 91.**Gibbons P, Flores de Arevalo H, Monico M. Assessment of the factor structure and reliability of the 28 item version of the general health questionnaire (GHQ-28) in El Salvador. *Int J Clin and Health Psychol*. 2004;4(2):389-398.
- 92.**Koc R, Erdemoglu AK. Validity and reliability of the Turkish Self-administered Leeds Assessment of Neuropathic Symptoms and Signs (S-LANSS) questionnaire. *Pain Med*. Jul 2010;11(7):1107-1114.
- 93.**Schneider G, Jull G, Thomas K, et al. Intra-rater and inter-rater reliability of select clinical tests in patients referred for diagnostic facet joint blocks in the cervical spine. *Arch Phys Med Rehabil*. 2013;In press.
- 94.**Norman G, Streiner D. *Biostatistics: The bare essentials*. 2 ed. Hamilton: BC Decker Inc; 2000.
- 95.**Kleinbaum D, Klein M. *Logistic Regression: A self-learning text*. 2nd. ed. New York, NY: Springer-Verlag; 2002.
- 96.**Fritz J, Wainner R. Examining diagnostic tests: an evidence-based perspective. *Phys Ther*. 2001;81(9):1546-1564.
- 97.**Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid?

Evidence-Based Medicine Working Group. *JAMA : the journal of the American Medical Association*. Feb 2 1994;271(5):389-391.

- 98.** Siegenthaler A, Eichenberger U, Schmidlin K, Arendt-Nielsen L, Curatolo M. What does local tenderness say about the origin of pain? An investigation of cervical zygapophysial joint pain. *Pain Med*. 2010;110(3):923-927.
- 99.** Schellingerhout JM, Heymans MW, de Vet HC, Koes BW, Verhagen AP. Categorizing continuous variables resulted in different predictors in a prognostic model for nonspecific neck pain. *J Clin Epidemiol*. Aug 2009;62(8):868-874.
- 100.** Schneider GM, Smith AD, Hooper A, et al. Minimizing the source of nociception and its concurrent effect on sensory hypersensitivity: an exploratory study in chronic whiplash patients. *BMC Musculoskelet Disord*. 2010;11:29.
- 101.** Manchikanti L, Cash K, Pampati V, Fellows B. Influence of psychological variables on the diagnosis of facet joint involvement in chronic spinal pain. *Pain Physician*. 2008;11:145-160.
- 102.** Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Ann Intern Med*. Feb 7 2006;144(3):201-209.
- 103.** Childs JD, Cleland JA. Development and application of clinical prediction rules to improve decision making in physical therapist practice. *Phys Ther*. Jan 2006;86(1):122-131.

## Appendix A Self report measures

### Pain Catastrophizing Scale

Sullivan MJL, Bishop S, Pivik J. (1995)

---

<b>Name:</b>	<b>Age:</b>	<b>Gender:</b>	<b>Date:</b>
-----	-----	<input type="checkbox"/> Male <input type="checkbox"/> Female	-----

---

Everyone experiences painful situations at some point in their lives. Such experiences may include headaches, tooth pain, joint or muscle pain. People are often exposed to situations that may cause pain such as illness, injury, dental procedures or surgery.

Instructions:

*We are interested in the types of thoughts and feelings that you have when you are in pain. Listed below are thirteen statements describing different thoughts and feelings that may be associated with pain. Using the following scale, please indicate the degree to which you have these thoughts and feelings when you are experiencing pain.*

RATING	0	1	2	3	4
<b>MEANING</b>	Not at all	To a slight degree	To a moderate degree	To a great degree	All the time

***When I'm in pain ...***

Number	Statement	Rating
1	I worry all the time about whether the pain will end.	
2	I feel I can't go on.	
3	It's terrible and I think it's never going to get any better	
4	It's awful and I feel that it overwhelms me.	
5	I feel I can't stand it anymore	
6	I become afraid that the pain will get worse.	
7	I keep thinking of other painful events	
8	I anxiously want the pain to go away	
9	I can't seem to keep it out of my mind	
10	I keep thinking about how much it hurts.	
11	I keep thinking about how badly I want the pain to stop	
12	There's nothing I can do to reduce the intensity of the pain	
13	I wonder whether something serious may happen.	

Copyright 1995 Michael J.L. Sullivan. Reproduced with permission.  
Source: Sullivan MJL, Bishop S, Pivik J. The pain catastrophizing scale: development and validation. *Psychol Assess*, 1995, 7: 524-532

## Neck Disability Index

This questionnaire has been designed to give the physiotherapist information as to how your neck pain has affected your ability to manage in everyday life. Please answer every section and mark in each section only the **ONE** box which applies to you. We realize you may consider that two of the statements in any one section relate to you, but please just mark the box which most closely describes your problem over the last 24 hours.

<p><b>SECTION 1 – PAIN INTENSITY</b></p> <p><input type="checkbox"/> I have no pain at the moment.</p> <p><input type="checkbox"/> The pain is very mild at the moment.</p> <p><input type="checkbox"/> The pain is moderate at the moment.</p> <p><input type="checkbox"/> The pain is fairly severe at the moment.</p> <p><input type="checkbox"/> The pain is very severe at the moment.</p> <p><input type="checkbox"/> The pain is the worst imaginable at the moment.</p>	<p><b>SECTION 6 – CONCENTRATION</b></p> <p><input type="checkbox"/> I can concentrate fully when I want to with no difficulty.</p> <p><input type="checkbox"/> I can concentrate fully when I want to with slight difficulty.</p> <p><input type="checkbox"/> I have a fair degree of difficulty in concentrating when I want to.</p> <p><input type="checkbox"/> I have a lot of difficulty in concentrating when I want to.</p> <p><input type="checkbox"/> I have a great deal of difficulty in concentrating when I want to.</p> <p><input type="checkbox"/> I cannot concentrate at all.</p>
<p><b>SECTION 2 – PERSON CARE (Washing, Dressing, etc.)</b></p> <p><input type="checkbox"/> I can look after myself normally without causing extra pain.</p> <p><input type="checkbox"/> I can look after myself normally but it causes extra pain.</p> <p><input type="checkbox"/> It is painful to look after myself and I am slow and careful.</p> <p><input type="checkbox"/> I need some help but manage most of my personal care.</p> <p><input type="checkbox"/> I need help every day in most aspects of self care.</p> <p><input type="checkbox"/> I do not get dressed, I wash with difficulty and stay in bed.</p>	<p><b>SECTION 7 – WORK</b></p> <p><input type="checkbox"/> I can do as much work as I want to.</p> <p><input type="checkbox"/> I can only do my usual work, but no more.</p> <p><input type="checkbox"/> I can do most of my usual work, but no more.</p> <p><input type="checkbox"/> I cannot do my usual work.</p> <p><input type="checkbox"/> I can hardly do any work at all.</p> <p><input type="checkbox"/> I can't do any work at all.</p>
<p><b>SECTION 3 – LIFTING</b></p> <p><input type="checkbox"/> I can lift heavy weights without extra pain.</p> <p><input type="checkbox"/> I can lift heavy weights but it gives extra pain.</p> <p><input type="checkbox"/> Pain prevents me from lifting heavy weights off the floor, but I can manage if they are conveniently positioned, for example on a table.</p> <p><input type="checkbox"/> Pain prevents me from lifting heavy weights, but I can manage light to medium weights if they are conveniently positioned.</p> <p><input type="checkbox"/> I can lift very light weights.</p> <p><input type="checkbox"/> I cannot lift or carry anything at all.</p>	<p><b>SECTION 8 – DRIVING</b></p> <p><input type="checkbox"/> I can drive my car without any neck pain.</p> <p><input type="checkbox"/> I can drive my car as long as I want with slight pain in my neck.</p> <p><input type="checkbox"/> I can drive my car as long as I want with moderate pain in my neck.</p> <p><input type="checkbox"/> I can't drive my car as long as I want because of moderate pain in my neck.</p> <p><input type="checkbox"/> I can hardly drive at all because of severe pain in my neck.</p> <p><input type="checkbox"/> I can't drive my car at all.</p>
<p><b>SECTION 4 – READING</b></p> <p><input type="checkbox"/> I can read as much as I want to with no pain in my neck.</p> <p><input type="checkbox"/> I can read as much as I want to with slight pain in my neck.</p> <p><input type="checkbox"/> I can read as much as I want to with moderate pain in my neck.</p> <p><input type="checkbox"/> I can't read as much as I want to because of moderate pain in my neck.</p> <p><input type="checkbox"/> I can hardly read at all because of severe pain in my neck.</p> <p><input type="checkbox"/> I cannot read at all.</p>	<p><b>SECTION 9 – SLEEPING</b></p> <p><input type="checkbox"/> I have no trouble sleeping.</p> <p><input type="checkbox"/> My sleep is slightly disturbed (less than 1 hour sleepless).</p> <p><input type="checkbox"/> My sleep is mildly disturbed (1-2 hrs. sleepless).</p> <p><input type="checkbox"/> My sleep is moderately disturbed (2-3 hrs. sleepless).</p> <p><input type="checkbox"/> My sleep is greatly disturbed (3-5 hrs. sleepless).</p> <p><input type="checkbox"/> My sleep is completely disturbed (5-7 hrs. sleepless).</p>
<p><b>SECTION 5 – HEADACHES</b></p> <p><input type="checkbox"/> I have no headaches at all.</p> <p><input type="checkbox"/> I have slight headaches which come in-frequently.</p> <p><input type="checkbox"/> I have moderate headaches which come in-frequently.</p> <p><input type="checkbox"/> I have moderate headaches which come frequently.</p> <p><input type="checkbox"/> I have severe headaches which come frequently.</p> <p><input type="checkbox"/> I have headaches almost all the time.</p>	<p><b>SECTION 10 – RECREATION</b></p> <p><input type="checkbox"/> I am able to engage in all my recreation activities with no neck pain at all.</p> <p><input type="checkbox"/> I am able to engage in all my recreation activities, with some pain in my neck.</p> <p><input type="checkbox"/> I am able to engage in most, but not all of my usual recreation activities because of my neck pain.</p> <p><input type="checkbox"/> I am able to engage in a few of my usual recreation activities because of pain in my neck.</p> <p><input type="checkbox"/> I can hardly do any recreation activities because of pain in my neck.</p> <p><input type="checkbox"/> I can't do any recreation activities at all.</p>

<b>TOTAL SCORE = _____/50</b>
-------------------------------

## 28-Item General Health Questionnaire

We should like to know if you have had any medical complaints, and how your health has been in general, over the past few weeks. Please answer ALL the questions on the following pages simply by underlining the answer which you think most nearly applies to you. Remember that we want to know about present and recent complaints, not those that you had in the past. It is important that you try to answer ALL the questions. Thank you very much for your co-operation.

A1. Have you recently been feeling perfectly well and in good health?

- better than usual
- same as usual
- worse than usual
- much worse than usual

A2. Have you recently been feeling in need of some medicine to pick you up?

- not at all
- no more than usual
- rather more than usual
- much more than usual

A3. Have you recently been feeling run down and out of sorts?

- not at all
- no more than usual
- rather more than usual
- much more than usual

A4. Have you recently felt that you are ill?

- not at all
- no more than usual
- rather more than usual
- much more than usual

A5. Have you recently been getting any pains in your head?

- not at all
- no more than usual
- rather more than usual
- much more than usual

A6. Have you recently been getting a feeling of tightness or pressure in your head?

- not at all
- no more than usual
- rather more than usual
- much more than usual

A7. Have you recently been having hot or cold spells?

- not at all
- no more than usual
- rather more than usual
- much more than usual

B1. Have you recently lost much sleep over worry?

- not at all
- no more than usual
- rather more than usual
- much more than usual

B2. Have you recently had difficulty staying asleep once you are off?

- not at all
- no more than usual
- rather more than usual
- much more than usual

B3. Have you recently felt constantly under strain?

- not at all
- no more than usual
- rather more than usual
- much more than usual

B4. Have you recently been getting edgy and bad-tempered?

- not at all
- no more than usual
- rather more than usual
- much more than usual

B5. Have you recently been getting scared or panicky for no good reason?

- not at all
- no more than usual
- rather more than usual
- much more than usual

B6. Have you recently found everything getting on top of you?

- not at all
- no more than usual
- rather more than usual
- much more than usual

B7. Have you recently been feeling nervous and uptight all the time?

- not at all
- no more than usual
- rather more than usual
- much more than usual

C1. Have you recently been managing to keep yourself busy and occupied?

- more so than usual
- same as usual
- rather less than usual
- much less than usual

C2. Have you recently been taking longer over the things you do?

- quicker than usual
- same as usual
- longer than usual
- much longer than usual

C3. Have you recently felt on the whole you were doing things well?

- better than usual
- about the same
- less well than usual
- much less well

C4. Have you recently been satisfied with the way you've carried out your task?

- more satisfied
- about the same as usual
- less satisfied than usual
- much less satisfied

C5. Have you recently felt that you are playing a useful part in things?

- more so than usual
- same as usual
- less useful than usual
- much less useful

C6. Have you recently felt capable of making decisions about things?

- more so than usual
- same as usual
- less so than usual
- much less capable



C7. Have you recently been able to enjoy your normal day-to-day activities?

- more so than usual
- same as usual
- rather less than usual
- much less than usual

D1. Have you recently been thinking of yourself as a worthless person?

- not at all
- no more than usual
- rather more than usual
- much more than usual

D2. Have you recently felt that life is entirely hopeless?

- not at all
- no more than usual
- rather more than usual
- much more than usual

D3. Have you recently felt that life isn't worth living?

- not at all
- no more than usual
- rather more than usual
- much more than usual

D4. Have you recently thought of the possibility that you might do away with yourself?

- definitely not
- I don't think so
- has crossed my mind
- definitely has

D5. Have you recently found at times you couldn't do anything because your nerves were too bad?

- not at all
- no more than usual
- rather more than usual
- much more than usual

D6. Have you recently found yourself wishing you were dead and away from it all?

- not at all
- no more than usual
- rather more than usual
- much more than usual

D7. Have you recently found that the idea of taking your own life kept coming into your mind?

- definitely not
- I don't think so
- has crossed my mind
- definitely has

## S-LANSS

1. **In the area where you have pain, do you also have 'pins and needles', tingling or prickling sensations?**
  - a) NO – I don't get these sensations
  - b) YES – I get these sensations often
  
2. **Does the painful area change colour (perhaps looks mottled or more red) when the pain is particularly bad?**
  - a) NO – The pain does not affect the colour of my skin
  - b) YES – I have noticed that the pain does make my skin look different from normal
  
3. **Does your pain make the affected skin abnormally sensitive to touch? Getting unpleasant sensations or pain when lightly stroking the skin might describe this.**
  - a) NO - The pain does not make my skin in that area abnormally sensitive to touch
  - b) YES – My skin in that area is particularly sensitive to touch
  
4. **Does your pain come on suddenly and in bursts for no apparent reason when you are completely still? Words like 'electric shocks', jumping and bursting might describe this.**
  - a) NO – My pain doesn't really feel like this
  - b) YES – I get these sensations often
  
5. **In the area where you have pain, does your skin feel unusually hot like a burning pain?**
  - a) NO – I don't have burning pain
  - b) YES – I get burning pain often
  
6. **Gently rub the painful area with your index finger and then rub a non-painful area (for example, an area of skin further away or on the opposite side from the painful area). How does this rubbing feel in the painful area?**
  - a) The painful area feels no different from the non-painful area
  - b) I feel discomfort, like pins and needles, tingling or burning in the painful area that is different from the non-painful area
  
7. **Gently press on the painful area with your finger tip then gently press in the same way onto a non-painful area (the same non-painful area that you chose in the last question). How does this feel in the painful area?**
  - a) The painful area does not feel different from the non-painful area
  - b) I feel numbness or tenderness in the painful area that is different from the non-painful area

## Appendix B Clinical tests

**Procedures used for the measurement of cervical ROM:** A universal inclinometer was used to measure flexion, extension, and side flexion. Use of a standard dual-arm goniometer was used to measure rotation.

### **Starting Position:**

Before taking any measurements all patients were instructed to “sit upright” and to keep their eyes focused “straight ahead.” Prior to movement testing, patients reported their current level of pain symptoms on a numeric pain rating scale and were instructed that these symptoms served as a baseline.

### **Neck Flexion and Extension:**

For neck flexion, the inclinometer was placed on the top of the patient’s head aligned with the external auditory meatus and then zeroed. The patient was asked to flex their head forward as far as possible, bringing the chin to the chest. The amount of neck flexion was recorded from the inclinometer. For extension ROM, the inclinometer was positioned in the same manner, and the patient was asked to extend their neck backwards as far as possible. The amount of neck extension was recorded from the inclinometer.

### **Neck Side-flexion:**

The inclinometer was positioned in the frontal plane on the apex of the patient’s head in alignment with the external auditory meatus. To measure right side-flexion, the patient was asked to move the right ear to the right shoulder. The amount of side-flexion was recorded from the inclinometer. The opposite was performed to measure left side-flexion. Care was taken to avoid concomitant rotation or flexion with the side-flexion movement.

### **Neck Rotation:**

Rotation was measured with a standard dual-arm goniometer. The patient was seated, looking directly forward with the neck in a neutral position. The fulcrum of the goniometer was placed over the top of the head with the stationary arm aligned with the acromion process, and the moveable arm bisecting the patient’s nose. The patient was asked to rotate their head and neck in each direction as far as possible.

### **Symptom Response:**

The effect of each movement on symptoms was noted and recorded as follows: 1) no effect, 2) increased symptoms, 3) decreased symptoms.

## **Manual Spinal Examination (MSE):**

### **Assessment:**

- 1) Patient will be positioned in prone with their cervical spine in a neutral position.
- 2) Starting with C2, palpate the spinous process with your thumb (the first bifurcated spinous process you can palpate in the upper cervical spine).
- 3) For the right C2 facet joint, move your thumb directly lateral to the right of the spinous process and you will contact the superior articular process of C2.
- 4) Apply a posterior-anterior directed force over the superior articular pillar of C2.
- 5) Repeat steps 1-4 for the facet joints (right and left) of C3-4 to 6-7.
- 6) Therapist assesses the quality (stiff, compressed, normal), quantity of motion (normal, abnormal), and end-feel (early capsular, spasm, normal = capsular) of the facet joint.
- 7) The therapist will categorize the displacement (quantity of motion) and tissue resistance (quality of motion) of the spinal segment qualitatively (normal, slight, moderate, or marked)
- 8) Therapist assesses whether this test provokes familiar pain.

### **Positive Test:**

- 1) Categorized as moderate or marked resistance to motion, and...
- 2) Familiar pain is provoked with the test ( $\geq 3/10$  on a NPRS).

### **Extension-Rotation Test (ER):**

#### **Assessment:**

- 1) Patient will be positioned in sitting and is asked to “sit upright” and “look straight ahead”.
- 2) Patient will be asked to, “extend their head and neck backward as far as possible”. When they get to the end of range, than they will be asked to, “bring their right ear to their right shoulder as far as possible”.
- 3) This test will be repeated to the left side.

#### **Positive Test:**

- 1) The test will be considered positive if the patient notes an increase in familiar pain symptoms ( $\geq 3/10$  on a NPRS) at the end of range of motion of extension and rotation.

## **Palpation presence of paraspinal tenderness (PST):**

### **Assessment:**

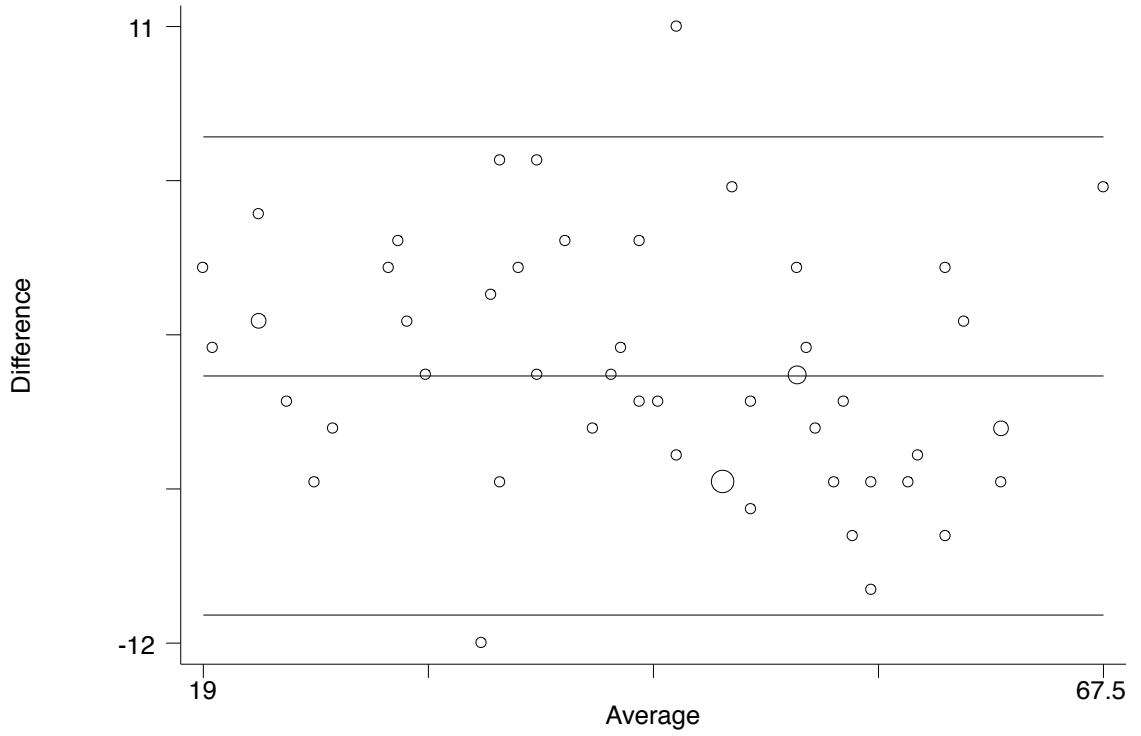
- 1) Patient will be positioned in prone with their cervical spine in the neutral position.
- 2) Therapist will apply a pressure of approximately 4 kg over the segmental muscle overlying the superior articular pillar of the facet joint (i.e.; for C2 – palpate the spinous process with your thumb; it is the first bifurcated spinous process you can palpate in the upper cervical spine; for the right C2 facet joint, move your thumb directly lateral to the right of the spinous process and you will contact the superior articular process of C2).
- 3) This test will be repeated for the left side.
- 4) This test will be repeated for the spinal levels C3-7.

### **Positive Test:**

- 1) Patient reports familiar pain ( $\geq 3/10$  on a NPRS) at the spinal level being tested that is noticeably greater than either the opposing side or adjacent spinal levels.

Appendix C

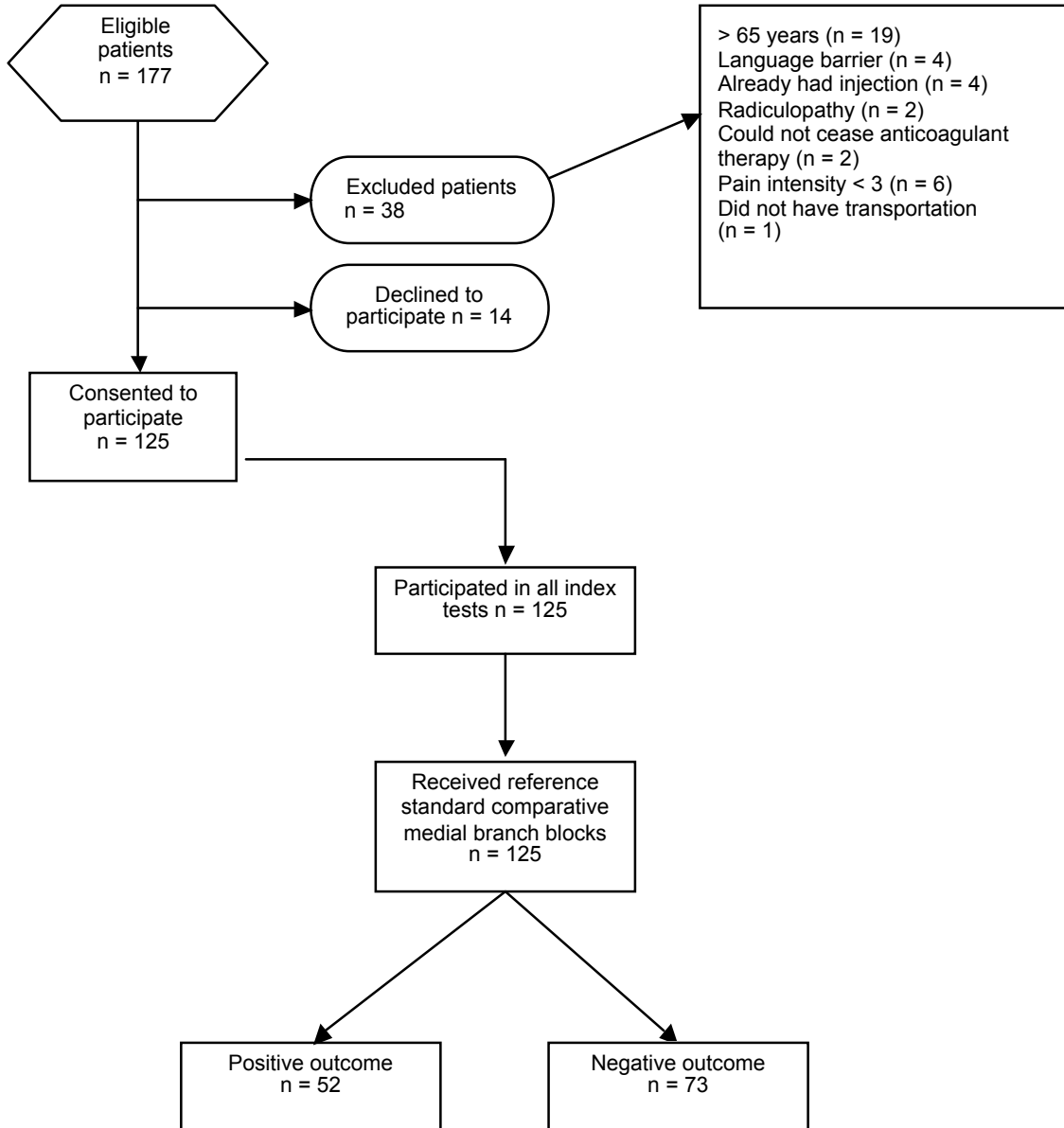
**Bland-Altman plot of the difference between raters for the measure of cervical spine flexion range of motion:**





Appendix D

Figure 1: Flow diagram of subject recruitment and participation



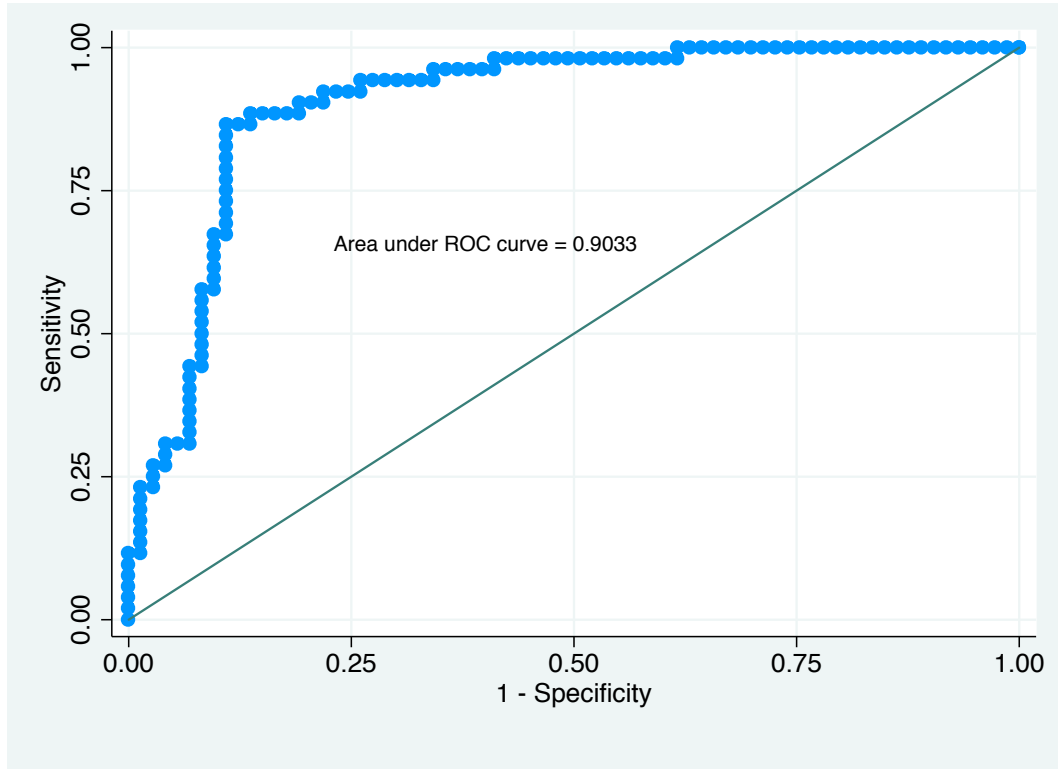
Appendix E

**Multivariate logistic regression model with primary exposure variables PST and ER, controlling for age, gender, baseline neck pain and disability, catastrophization, and psychological distress (p<0.001):**

<b>Predictor Variable</b>	<b>Coefficient</b>	<b>Standard Error</b>	<b>z-statistic</b>	<b>p-value</b>	<b>95% CI</b>
PST	3.71	0.71	5.23	<0.001	2.32-5.10
ER	1.50	0.62	2.43	0.015	0.29-2.71
Age	-0.02	0.03	-0.70	0.482	-0.07-0.03
Gender	-0.02	0.62	-0.04	0.969	-1.24-1.19
Baseline neck pain and disability	-0.06	0.05	-1.30	0.194	-0.15-0.03
Catastrophization	0.001	0.03	0.04	0.971	-0.06-0.07
Psychological distress	-0.04	0.03	-1.41	0.159	-0.10-0.02
Constant	-0.69	1.69	-0.41	0.685	-4.00-2.63

Appendix F

**ROC curve for the multivariate logistic regression model with primary exposure variables PST and ER, controlling for age, gender, baseline neck pain and disability, catastrophization, and psychological distress:**



Appendix G

**Multivariate logistic regression model with primary exposure variables MSE and ER, controlling for age, gender, baseline neck pain and disability, catastrophization, and psychological distress (p<0.001):**

<b>Predictor Variable</b>	<b>Coefficient</b>	<b>Standard Error</b>	<b>z-statistic</b>	<b>p-value</b>	<b>95% CI</b>
MSE	3.35	0.65	5.17	<0.001	2.08-4.62
ER	1.38	0.69	2.30	0.021	0.20-2.55
Age	-0.02	0.03	-0.92	0.356	-0.07-0.03
Gender	-0.06	0.60	-0.10	0.919	-1.23-1.11
Baseline neck pain and disability	-0.05	0.04	-1.17	0.243	-0.14-0.04
Catastrophization	-0.01	0.03	-0.19	0.85	-0.07-0.06
Psychological distress	-0.04	0.03	-1.54	0.123	-0.10-0.01
Constant	-0.05	1.59	-0.03	0.974	-3.17-3.06

## Appendix H

**ROC curve for the multivariate logistic regression model with primary exposure variables MSE and ER, controlling for age, gender, baseline neck pain and disability, catastrophization, and psychological distress:**

