Is there preliminary value to a within- and/or between-session change for determining short-term outcomes of manual therapy on mechanical neck pain?

Chad Cook¹, Jessica Lawrence¹, Katelyn Michalak¹, Sidra Dhiraprasiddhi¹, Megan Donaldson¹, Shannon Petersen², Kenneth Learman³

¹Walsh University, USA, ²University of Des Moines, USA, ³Youngstown State University, USA

Objectives: The primary purpose of this study was to determine whether occurrences of within- and between-session changes were significantly associated with functional outcomes, pain, and self-report of recovery in patients at discharge who were treated with manual therapy for mechanical neck pain. A secondary purpose was to determine the extent of change needed for the within- and between-session change in association to function.

Methods: This secondary data analysis examined 56 patients who demonstrated a positive response to manual therapy during the initial assessment within a randomized controlled trial (RCT) that examined manual therapy techniques and a home exercise program (HEP). Within- and between-session findings were defined as 'changes in pain report during the initial session (within)' and 'changes in pain from baseline to 48-hours post initial assessment (between)'. Outcomes were analyzed for associations with the global rating of change (GROC), self-report activity scale (SRAS), and a 50% reduction of the neck disability index (NDI) by discharge at 96 hours.

Results: Findings indicate that within-session pain changes of 36.7% are strongly associated with a 50% change in NDI at 96 hours. Between-session changes in pain were associated with 50% change in NDI and a ≥ 3 -point change in GROC at 96 hours.

Conclusion: Both within- and between-session measures may be useful to predict success levels at 96 hours for NDI; however, between-session changes are more useful to predict success in GROC. Measures used during clinical examination may help guide clinicians in identification of candidates best suited for the treatment.

Keywords: Within-session change, Between-session change, Mechanical neck pain, NDI, GROC

Introduction

Neck pain is a significant financial burden for society.¹ In 2006, over 13 million patient visits were for neck-oriented problems in the United States, a total equating to 1% of *all* health care visits to hospitals and physician offices.² Conservative treatment methods are typically recommended and utilized for neck pain.³ Of the conservative options, manual therapy and exercise exhibit the most compelling evidence of benefit^{4,5} and are advocated by mono-disciplinary guidelines.⁶ Despite this, non-advocated treatment components⁶ such as electrotherapy stimulation (30.3%), corsets or braces (20.9%), massage (28.1%), ultrasound (27.3%), heat (57.0%), and cold (47.4%), which have unclear or little

benefit in the recovery of cervical/neck pain are frequently used in conservative care of patients with neck pain.⁷ Manual therapy and exercises were only received 36.8 and 52.5% of the time, respectively.⁷

The use of passive modalities may be associated with the inability of a clinician to identify appropriate candidates for manual therapy or exercise. Historically, clinicians have relied on good clinical reasoning skills to identify appropriate candidates and just recently, classification of patients into treatment subgroups has gained interest. Markers for subgroup classification have become a priority for research for multiple musculoskeletal conditions.⁸⁻¹⁰ Classifying patients into subgroups requires the use of a formal assessment strategy and clinical examination findings that have prognostic utility, and there are presently a number of statistical processes that can be used.

Correspondence to: Chad Cook, Walsh University, North Canton, OH, 44720, USA. Email: ccook@walsh.edu

Within- and between-session changes during the clinical examination are two indicators that have been used and advocated for over 40 years.^{11,12} Just recently, a number of studies^{13–18} have examined the prognostic capacity of these indicators and their contributions toward clinical reasoning.

Within- and between-session changes are improvements in a patient's symptoms that have occurred secondary to the initial examination or treatment. Changes in pain or range of motion (ROM) during a single visit are termed as within-session changes.¹¹ Changes in pain or ROM that carry over to a succeeding visit are known as between-session change.¹¹ Both of these parameters have been evaluated in studies on patients with neck^{13–15} and low back pain^{15,16} and shoulder¹⁷ and hip pain.¹⁸ Current evidence provides the strongest support for within- and between-session changes within the spine.^{13–16} There is evidence that within-session changes are predictive of between-session changes for patients with neck pain; however, to date within-session changes do not appear to be predictive of long-term functional improvement.¹³ Currently, within- and between-session changes are not well-supported in studies on the extremities.^{17,18}

Past studies have used changes in pain and ROM as markers of change for patients. Cook and colleagues¹⁶ suggested that a 2-point between-session change in pain was required for a meaningful prediction of long-term functional outcome of patients with mechanical low back pain. At present, there are neither studies that have quantified a minimal predictive change value for patients with neck pain, nor studies that have examined the within- or between-session effect on outcome measures such as the global rating of change (GRoC) or self-report activity scale (SRAS). Consequently, the aim of this study was to determine the predictive validity of change in pain for short-term (96 hour) improvements and different outcomes measures in a population with neck pain. The clinical significance of this study is to quantify the level of pain-related change needed for functional improvements and whether pain-related improvements are influential in different forms of outcomes measures.

Methods

Study design

This study was a secondary database analysis of a randomized clinical trial (RCT)¹⁹ involving 56 participants who received manual therapy interventions and home exercise program (HEP) for mechanical neck pain. Refer to [clinicaltrials.gov #NCT01750736](https://clinicaltrials.gov/ct2/show/study/NCT01750736)¹⁹ for further information regarding the primary study.

Participants

Enrollment criteria were designed to reflect the recommendations of the mono-disciplinary guidelines

for physical therapy.⁶ Participants were 18 years of age or older with neck pain of any duration. Inclusion criteria included the following: (1) primary complaint of unilateral neck pain, (2) complaints of neck motion limitations, (3) onset of symptoms that may be linked to a recent unguarded/awkward movement or position, (4) with or without associated upper extremity pain, (5) limited cervical ROM; restricted cervical extension and/or cervical rotation and/or cervical lateral flexion ROM, (6) neck pain reproduced at end ranges of active and passive motions, (7) restricted cervical and thoracic segmental mobility determined through passive accessory and passive physiological examination, and (8) neck and neck-related upper extremity pain reproduced with provocation of the involved cervical or upper thoracic segments.

Patients were excluded if red flags were noted or self-reported during the patient's medical screening including any one of the following signs or symptoms: tumor, fracture, metabolic diseases, rheumatoid arthritis, osteoporosis, prolonged history of steroid use, symptoms of vertebrobasilar insufficiency, current pregnancy, cervical spinal stenosis, or bilateral upper extremity symptoms, use of blood thinners, whiplash injury within the past 6 weeks, central nervous system involvement such as hyperreflexia, sensory disturbances in the hand, intrinsic muscle wasting of the hands, unsteadiness during walking, nystagmus, loss of visual acuity, impaired sensation of the face, altered taste, or exhibited pathological reflexes.

Further exclusion included two or more positive neurologic signs consistent with nerve root compression, including any two of the following: muscle weakness involving a major muscle group of the upper extremity, diminished upper extremity muscle stretch reflex, diminished or absent sensation to pinprick in any upper extremity dermatome, prior surgery to the neck or thoracic spine, report of workers compensation or pending legal action regarding their neck pain, insufficient English language skills to complete all questionnaires, or inability to comply with treatment and follow-up schedule.

Study treatment

The study treatment and data collection occurred at two academic locations in Northeast Ohio and one academic institution in Central Iowa. The in-clinic interventions were performed on the *first day only*, whereas all subsequent interventions were provided as HEP and were self-administered by the patient. The study only looked at 96 hours of outcomes (initial treatment + 4 days of exercises). Examination and treatment were provided by a fellowship trained manual physical therapist of the American Academy of Orthopedic Manual Physical Therapists (AAOMPT).

Targeted interventions included the choices of thrust and non-thrust manipulation to the cervical spine, and when deemed appropriate, muscle energy techniques, stretching, or other manual therapy-oriented procedures. Each patient was provided a HEP of either segment-specific exercises with use of a towel or strap to target specific segments/region of the neck (self-mobilization) with active movement of the desired neck motion or general ROM exercises. The general ROM exercises emphasized all motions including cervical rotation, lateral flexion, extension, and flexion exercises. The segment-specific exercises (self-mobilization) only performed the primary desired motion with segmental/regional isolation as prescribed based on the examination. The primary study associated with this trial showed no significant differences in outcomes between the two exercise interventions within the study population.

Outcomes measures

All outcome measures were administered at the initial evaluation (baseline) and at the follow-up examination 48 and 96 hours post baseline measurement.

Numeric pain rating scale

Previous studies have shown the numeric pain rating scale (NPRS) to be reliable and valid to measure pain intensity.^{20–23} Patients rated their current level of pain and their worst and least amount of pain experienced during the previous 24 hours. This 11-point scale is anchored on the left with the phrase ‘No Pain’ and on the right with the phrase ‘Worst Possible Pain’.

Global rating of change

Different forms of global rating of change (GROC) scales have been used as primary outcome measures or anchor measures in validation studies. Scales items vary significantly and range from 3 item, 7 item, and 15 item tools.²⁴ Despite the variations in number of items, the associated verbal representations of each scale allows patients to identify whether they have improved or worsened with qualifiers for each level of improvement such as minimally or substantially. For this study, we used a 15-point global rating scale ranging from -7 (a very great deal worse) to 0 (about the same) to $+7$ (a very great deal better).²⁴ Threshold scores (cut points) have ranged from values of 1 , 2 , >4 , and 5 or greater. Others²⁵ have suggested that since multiple numeric scales exist, the language should be considered to standardize each scale. For our study, we selected ‘somewhat’ as the equivalent of ‘minimally improved’ in order to reflect the language in the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials recommendations.²⁵

Neck disability index

This questionnaire measures pain-related disability associated with activities of daily living in people with

neck pain.²⁶ The neck disability index (NDI) is easy to complete and score; each item is recorded on a 6-point scale with a maximum score of 5, resulting in a maximum potential score of 50.²⁷ This score can be calculated as a percentage, with higher scores indicating higher levels of disability.^{26,27} In this study, NDI scores were reported as absolute values. Content and construct validity and reliability of the NDI have been previously shown in patients with neck pain.^{26,28} Cleland *et al.*²⁹ regarded the NDI as an appropriate tool to be used by researchers. The NDI has been used by other researchers to evaluate the effect of treatments on the patient’s perceived level of functioning and disability.³⁰

Self-report activity scale

This survey is a unique 15-point scale based on the GROC form that allows the participant to rate his/her perception of the change in activity level, which is a result of the treatment provided. A similar measurement tool of this scale has been used to capture self-report of overall and functional change in past clinical trials.³¹ The SRAS is a novel instrument created for the use of the parent randomized controlled trial (RCT).

Within-session changes

Within-session changes were captured at the initial visit. Pain was assessed prior to and immediately following a single treatment session of mobilization or manipulation of a symptomatic segment. Within-session scores were calculated using the following formula: $((\text{baseline pain} - \text{immediate post session pain}) / \text{baseline pain}) \times 100\%$.

Between-session changes

Between-session scores were captured from the initial visit to 48-hours post initial evaluation. Final outcomes were captured at 96 hours only, providing a 4-day outcome measure. As the pain values at baseline varied significantly, change scores were converted into percentage change from baseline. Between-session scores were calculated using the following formula: $((\text{baseline pain} - \text{second session pain}) / \text{baseline pain}) \times 100\%$. This allowed a capture of change on the NPRS over a period of two full visits. Within- or between-session percentage change scores could range from -100 to 100% , depending on the baseline score and the subsequent follow up score.

Dichotomized outcomes measures of success

The GROC, the NDI, and the SRAS were dichotomized and used as distinct measures of success. The GROC was dichotomized to ≥ 3 as ‘success’ whereas <3 was equal to a lack of success based on previous suggestions.²⁵ A 50% change from baseline or higher was considered a ‘success’ for the NDI, based on previous work using the Oswestry Disability Index in

patients with lumbar spine related problems.³² The SRAS was a novel measure for this study, with ≥ 3 considered 'success' based on our own cut point selection.

Data analysis

Data analysis included descriptive and inferential statistics using SPSS version 20.1. To determine the discrimination of percentage change in pain among within- and between-session changes, we used a receiver operating characteristic (ROC) curve for each of the dichotomized outcome measures of success. The ROC represents a graphical plot of the true- versus false-positive rates of a finding, when the outcome variable is dichotomous and the predictor variable is continuous. Numeric pain rating scale values were change scores during the first session (which reflected within-session changes) and between the baseline and follow up at 48 hours (which reflected between-session changes).

The area under the curve (AUC)³³ was used to define the extent of the relationship between the dependent and predictor variable, with values closer to 1 suggesting near perfect findings and values closer to 0.5 suggesting no relationship or value in the independent variable. The model required statistical significance ($P \leq 0.05$) to determine relevant findings.

If values were significant within the ROC curve, we calculated diagnostic accuracy statistics, including sensitivity, specificity, and positive and negative likelihood ratios, as well as associated 95% confidence intervals (95% CI). Post-test probability was analyzed by looking at the prevalence of a positive finding with the outcome and multiplying that prevalence by the positive and negative likelihood ratio. Lastly, a binomial logistic regression analysis was used to determine the overall effect of exceeding the threshold value. For all analyses, a P -value ≤ 0.05 was considered significant.

Results

All analyses were performed on 56 patients who met the inclusion criteria of a baseline NPRS score of greater than or equal to 2/10. Figure 1 outlines the process of study sample selection, data collection, and secondary data analysis. Table 1 outlines the descriptive characteristics of the patients for both exercise groups and the aggregate data for the sample with the results of the statistical analyses.

In detecting pertinent within-session percentage changes, only the dichotomized outcome measure of 50% change on the NDI produced a significant finding. A 36.7% change in pain during a within-session visit was strongly associated with a 50% change in the NDI at 96 hours (AUC=0.665, $P=0.04$). Between-session NPRS changes were also associated with the dichotomized outcome measures,

50% change on the NDI at 96 hours (AUC=0.864, $P<0.01$) and a ≥ 3 -point change on the GROC at 96 hours (AUC=0.761, $P<0.01$). A change of 36.7% from baseline was also the most robust value associated with the two anchors (Table 2).

The sensitivity, specificity, and positive and negative likelihood ratios for the between-session changes for the NDI and GROC and the within-session changes on the NDI were analyzed and 36.7% change in pain was more sensitive than specific for a 50% change in the NDI, whereas the pain change associated with the GROC finding was more specific. The post-test probability of a negative outcome with no presence of a within-session and between-session change of the NDI was 23.1 and 13.1% change, respectively, suggesting that the tool is moderately useful in identifying those who are unlikely to benefit from the single manual therapy approach (Table 3).

A binary logistic regression analysis of significant variables from the ROC was conducted to identify odds ratios and confidence intervals. Patients who experienced a 36.7%, between-sessions change in pain were 21.7 times more likely to report an improvement of $\geq 50\%$ on the NDI ($P<0.01$) than those who did not achieve a 36.7% change in pain. Those with a 36.7%, between-sessions change in pain were also 7.3 times more likely to report an improvement of ≥ 3 on the GROC ($P<0.01$) than those who did not achieve a 36.7% change in pain. Lastly, a within-session change of 36.7% were 5.7 times as likely to report an improvement of $\geq 50\%$ on the NDI ($P<0.01$) than those who did not achieve a 36.7% change in pain (Table 4).

Discussion

The purpose of this secondary data analysis study was to determine the predictive validity of change in pain toward short-term (96 hour) improvements and disparate outcomes measures on a patient population with mechanical neck pain. Our inclusion criteria were designed to reflect the recommendations of the mono-disciplinary guidelines⁶ for physical therapy and to allow replication of sampling in future studies. Change in pain was assessed by looking at within-session changes (changes within the same visit) and between-session changes (changes from the first to the next visit-in this study, 48 hours later). Both measures were able to predict pre-determined success levels at 96 hours for NDI, whereas only between-session changes predicted success in the GROC score. In all three measures, a pain change score of 36.7% was the most robust change associated with the targeted outcomes measures. Of particular interest is the fact that only one manual therapy intervention was provided (at the first visit), but the outcomes were still positive in nearly 50% of all patients.

Figure 1 Flow and selection process within the study. RCT: randomized clinical trial; NPRS: numeric pain rating scale; NDI: neck disability index.

Table 1 The descriptive characteristics of the primary study¹⁹ patients receiving manual therapy. Augmented and non-augmented exercise groups combined (N=56)

Variable	Mean (SD); minimum–maximum
Age (years)	33.9 (14.8); 18–72
Gender	22=Male 34=Female
Race	55=White 0=Black 0=Hispanic 1=Asian 0=Other
Height (m)	1.68 (0.11); 1.42–2.01
Weight (kg)	71.4 (19.96); 44.45–117.93
Body mass index (kg/m ²)	25.0 (4.3); 16.8–35.2
Duration of symptoms (weeks)	129.6 (264.4); 0.5–1664
NDI (absolute score)	12.6 (6.5); 3–44
NPRS	3.3 (1.4); 2–6.8

NDI: neck disability index; NPRS: numeric pain rating scale (baseline).

Previous literature has investigated the within- and between-session effects toward outcomes at the cervical spine. Tuttle¹⁵ reported that within-session changes may predict between-session changes, and are related to changes in impairments such as ROM. In contrast to our findings, Tuttle¹⁵ did not report improvement in functional changes after within- and between-session changes. Differences between the two studies include sample size (our study sample, although only 56, was nearly double that of Tuttle's 29 patients), sample characteristics (our population was less impaired than Tuttle's), and duration of care. Tuttle treated patients up to six visits over a 2 weeks span. Within our design, only one formal intervention was provided (on the first day), although each patient was instructed to perform a HEP within the 96 hour span of the study.

Table 2 ROC analysis for anchor measures of 50% change in NDI, 3-point score or greater on the GRoC, and 3-point score of greater on the SRAS

Variables	AUC	P-value	Recommended NPRS threshold score
Within-session change			
50% change in NDI	0.665	$P=0.04$	>36.7% change in pain
≥3 on the GRoC	0.611	$P=0.16$	N/A
≥3 on the SRAS	0.501	$P=0.99$	N/A
Between-session change			
50% change in NDI	0.864	$P<0.01$	>36.7% change in pain
≥3 on the GRoC	0.761	$P<0.01$	>36.7% change in pain
≥3 on the SRAS	0.548	$P=0.62$	N/A

ROC: receiver operator characteristics curve; NDI: neck disability index; GRoC: global rating of change; SRAS: self-report activity scale; NPRS: numeric pain rating scale; N/A: not applicable; AUC: area under the curve.

A meaningful element of our study was the finding that a 36.7% change in pain for within- and between-session changes was associated with a 50% reduction in the NDI and an improvement of ≥ 3 on the GRoC. Other studies^{13–18} used point estimates to analyze within- and between-session changes. O'Halloran *et al.*³⁴ suggested that percent change is likely to provide more utility among different patient severity levels and is able to capture a positive outcome in situations where ceiling or floor effects may be present. Indeed, the individuals in our study demonstrated lower levels of baseline severity (NDI=12.6 [SD=6.5] and NPRS=3.3 [SD=1.4]) and if we used a point value as a change score as others have in the past, those with NPRSs of 2 or 3 could have potentially been omitted as 'successes'. We feel that the findings associated with the single 36.7% change in pain may have utility for clinical practice, as it could provide a framework for clinicians to guide decision making that is likely transferable to multiple populations. We also find it coincidental that the value associated with a meaningful disability improvement is consistent with previous research examining a clinically meaningful change in pain identified on the NPRS. Farrar *et al.*³⁵ found that a 30% change score was indicative of clinically meaningful change in pain and the change was consistent across baseline scores. Numerous other investigators have suggested that changes of 33–50% are benchmark standards for meaningful pain reduction.^{35–38}

The use of percent change scores for pain rating is not novel. It has been previously suggested that percent change scores should be used to provide a definition of a clinically meaningful change. This is because minimum clinically important differences for pain ratings is dependent on mean baseline score of the sample being studied, suggesting that a percent change would be a more valid concept. Farrar *et al.*³⁶ suggested that most patients actually think of pain improvement as a percent change rather than a raw pain score. Thus, contextually the use of a percent change coincides with patient concepts. Others have supported this concept as Price and colleagues³⁹ recommended using the pain scale as a ratio scale.

One problem with our findings is the application of the 36.7% change score found in clinical practice. The percent change observed is clearly a mean effect of a larger sample and most individuals will struggle with the idea of identifying a 36.7% change in their pain level. Previous investigators have used the 50% benchmark cut-off for change because it is easy for both patients and clinicians to understand and apply; however, the value is high and may cause the clinician to under appreciate clinically meaningful improvements in some patients.³⁶ An additional problem may arise since it has been previously found that percent change scores are not always normally distributed which can result in a situation where the mean value does not accurately reflect central tendency.⁴⁰ Regardless of these potential problems, the remarkable consistency

Table 3 Sensitivity, specificity, positive and negative likelihood ratios and post-test probability findings of significant variables

Variable	Sensitivity (95% CI)	Specificity (95% CI)	Positive likelihood ratio (95% CI)	Negative likelihood ratio (95% CI)	Post-test probability of a positive finding	Post-test probability of a negative finding
Within-session change of pain for the NDI	75.9 (57.9, 88.0)	65.6 (52.0, 75.4)	2.2 (1.2, 3.6)	0.4 (0.2, 0.8)	62.3%	23.1%
Between-session change of pain for the NDI	83.3 (67.1, 93.6)	81.3 (69.1, 88.9)	4.4 (2.2, 8.4)	0.2 (0.1, 0.5)	76.8	13.1%
Between-session change of pain for the GRoC	65.6 (57.9, 74.6)	79.2 (62.2, 91.1)	3.2 (1.4, 8.4)	0.4 (0.3, 0.8)	80.9%	34.7%

CI: confidence interval; NDI: neck disability index; GRoC: global rating of change. Pre-test probability of success for NDI=42.9%. Pre-test probability for success of GRoC=57.1%.

Table 4 Binary logistic regression analysis of significant variables from the ROC curves

Variable	Odds ratio (95% CI)	P value
Within-session change of pain for the NDI	5.7 (1.7, 18.5)	<0.01
Between-session change of pain for the NDI	21.7 (5.4, 87.2)	<0.01
Between-session change of pain for the GroC	7.3 (2.1, 24.7)	<0.01

CI: confidence interval; NDI: neck disability index; GroC: global rating of change; ROC: receiver operating characteristic.

of the meaningful percent change scores reported in the literature supports the validity of the findings.

An assumption associated with the use of within- and between-session findings is that early changes reflective of the care provided are able to identify those who will succeed with that dedicated form or care; just as a lack of change is likely to predict that an individual is not going to succeed with that form of care. In our study, the dedicated form of care was manual therapy. The model is designed to be used early in the clinical examination as a guide to decision making. Tests and measures used early in a clinical examination should demonstrate acceptable sensitivity to appropriately capture a large portion of the population who may be candidates for the treatment. This is much different than a clinical prediction rule, which is normally highly specific, suggesting that tool is useful in confirming a targeted population (and should be used at the end of an examination).⁴¹ The within- and between-session findings are sensitive and do appear to fit the model assumptions. For example, only 13.1% of individuals who failed to demonstrate a between-session change in pain of 36.7% went on to achieve a 50% change on the NDI. Slightly above 23% of individuals who failed to demonstrate a within-session change in pain of 36.7% went on to achieve a 50% change on the NDI. We feel that these findings suggest promise when combining this approach to an approach that confirms items with a clinical prediction rule.

Limitations

Limitations of the primary study include small sample size and lack of ethnic/racial diversity. Larger trials with longer and more comprehensive interventions are necessary to substantiate these preliminary findings. The sample size of 56 patients creates a wide confidence interval, thus creating difficulty in reaching clinical significance. Owing to the longevity of the patients' neck pain, outcome measures over a longer-term would be preferable. Furthermore, the lower levels of disability affect the subject's quality of life to a lesser extent; therefore, changes in disability will not be as significant when compared to a sample of patients with acute neck pain with higher reports of disability.

Conclusion

Within- and between-session changes appear useful in predicting selected functional and well-being oriented

outcome measures at 96 hours in a population of patients receiving manual therapy and self-applied exercises. These treatment decision-making elements may have clinical utility at the early phases of the clinical examination and may lend value, when combined with other decision-making tools, in determining best outcomes with specific interventions such as manual therapy. Future research is needed to investigate the clinical utility of within- and between-session change for use during clinical examination to guide decision making regarding intervention choices.

Disclaimer Statements

Contributors All authors contributed to the writing of the paper. Dr Cook performed all analyses and created the study idea.

Funding OPTA Cardon/AAOMPT.

Conflicts of interest None of the authors have a conflict of interest on this study. Dr Cook was funded through the OPTA. Dr Petersen and Dr Cook were funded through the AAOMPT/Cardon grant.

Ethics approval The study was approved by the Walsh University HSR, and the IRB's of Youngstown and Des Moines.

References

- 1 Cote P, Cassidy JD, Carroll L. The factors associated with neck pain and its related disability in the Saskatchewan population. *Spine (Phila Pa 1976)*. 2000;25:1109–17.
- 2 Driessen MT, Lin CW, van Tulder MW. Cost-effectiveness of conservative treatments for neck pain: a systematic review on economic evaluations. *Eur Spine J*. 2012;21:1441–50.
- 3 Irwin ZN, Hilibrand A, Gustavel M, McLain R, Shaffer W, Myers M, et al. Variation in surgical decision making for degenerative spinal disorders. Part II: cervical spine. *Spine (Phila Pa 1976)*. 2005;30:2214–9.
- 4 Miller J, Gross A, D'Sylva J, Burnie SJ, Goldsmith CH, Graham N, et al. Manual therapy and exercise for neck pain: a systematic review. *Man Ther*. 2010;15:334–54.
- 5 Gross A, Miller J, D'Sylva J, Burnie SJ, Goldsmith CH, Graham N, et al. Manipulation or mobilisation for neck pain: a Cochrane review. *Man Ther*. 2010;15:315–33.
- 6 Childs JD, Cleland JA, Elliott JM, Teyhen DS, Wainner RS, Whitman JM, et al. Neck pain: clinical practice guidelines linked to the International Classification of Functioning, Disability, and Health from the Orthopedic Section of the American Physical Therapy Association. *J Orthop Sports Phys Ther*. 2008;38:A1–34.
- 7 Goode AP, Freburger J, Carey T. Prevalence, practice patterns, and evidence for chronic neck pain. *Arthritis Care Res (Hoboken)*. 2010;62:1594–601.
- 8 Fritz JM, Brennan GP. Preliminary examination of a proposed treatment-based classification system for patients receiving physical therapy interventions for neck pain. *Phys Ther*. 2007;87:513–24.

- 9 Kay TM, Gross A, Goldsmith CH, Rutherford S, Voth S, Hoving JL, *et al.* Exercises for mechanical neck disorders. *Cochrane Database Syst Rev.* 2012;8:CD004250.
- 10 Sahrman SA. Diagnosis and treatment of movement impairment syndromes. St. Louis: Mosby, Inc; 2002.
- 11 Maitland GD, Hengeveld L, Banks K, English K. Maitland's vertebral manipulation, 6th edn. Oxford, UK: Butterworth-Heinemann; 2001.
- 12 Cook CE. Orthopedic manual therapy: an evidence-based approach. Upper Saddle River, NJ: Prentice Hall; 2007.
- 13 Tuttle N. Do changes within a manual therapy treatment session predict between-session changes for patients with cervical spine pain? *Aust J Physiother.* 2005;51:43–8.
- 14 Hahne AJ, Keating JL, Wilson SC. Do within-session changes in pain intensity and range of motion predict between-session changes in patients with low back pain? *Aust J Physiother.* 2004;50:17–23.
- 15 Tuttle N, Laasko L, Barrett R. Change in impairments in the first two treatments predicts outcome in impairments, but not in activity limitations, in subacute neck pain: an observational study. *Aust J Physiother.* 2006;52:281–5.
- 16 Cook CE, Showalter C, Kabbaz V, O'Halloran B. Can a within/between-session change in pain during reassessment predict outcome using a manual therapy intervention in patients with mechanical low back pain? *Man Ther.* 2012;17:325–9.
- 17 Garrison JC, Shanley E, Thigpen C, Hegedus E, Cook C. Between-session changes predict overall perception of improvement but not functional improvement in patients with shoulder impingement syndrome seen for physical therapy: an observational study. *Physiother Theory Pract.* 2011;27:137–45.
- 18 Wright AA, Abbott JH, Baxter D, Cook C. The ability of a sustained within-session finding of pain reduction during traction to dictate improved outcomes from a manual therapy approach on patients with osteoarthritis of the hip. *J Man Manip Ther.* 2010;18:166–72.
- 19 Peterson S, Cook C, Donaldson M, Hassen A, Learman K. Manual Therapy and Augmented Exercise for Neck Pain. *Clinicaltrials.gov.* Trial Registry #NCT01750736. Registered December 11, 2012.
- 20 Downie WW, Leatham PA, Rhind VM, Wright V, Branco JA, Anderson JA. Studies with pain rating scales. *Ann Rheum Dis.* 1978;37:378–81.
- 21 Jensen MP, Miller L, Fisher LD. Assessment of pain during medical procedures: a comparison of three scales. *Clin J Pain.* 1998;14:343–9.
- 22 Jensen MP, Turner JA, Romano JM. What is the maximum number of levels needed in pain intensity measurement? *Pain.* 1994;58:387–92.
- 23 Katz J, Melzack R. Measurement of pain. *Surg Clin North Am.* 1999;79:231–52.
- 24 Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials.* 1989;10:407–15.
- 25 Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, *et al.* Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain.* 2008;9:105–21.
- 26 Vernon H, Mior S. The neck disability index: a study of reliability and validity. *J Manipulative Physiol Ther.* 1991;14:409–15.
- 27 Finch E, Brooks D, Stratford P, Mayo N. Physical rehabilitation outcome measures: a guide to enhanced clinical decision making, 2nd edn. Ontario, BC: Decker; 2002.
- 28 Hoving JL, O'Leary EF, Niere KR, Green S, Buchbinder R. Validity of the neck disability index, Northwick Park neck pain questionnaire, and problem elicitation technique for measuring disability associated with whiplash-associated disorders. *Pain.* 2003;102:273–81.
- 29 Cleland JA, Fritz JM, Whitman JM, Palmer JA. The reliability and construct validity of the neck disability index and patient specific functional scale in patients with cervical radiculopathy. *Spine (Phila Pa 1976).* 2006;31:598–602.
- 30 Riddle DL, Stratford PW. Use of generic versus region-specific functional status measures on patients with cervical spine disorders. *Phys Ther.* 1998;78:951–63.
- 31 Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther.* 2009;17:163–70.
- 32 Fritz JM, Hebert J, Koppenhaver S, Parent E. Beyond minimally important change: defining a successful outcome of physical therapy for patients with low back pain. *Spine (Phila Pa 1976).* 2009;34:2803–9.
- 33 Akobeng AK. Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta Paediatr.* 2007;96:644–7.
- 34 O'Halloran B, Wright A, Cook CE. Criterion validation of the rate of recovery, a single alphanumeric measure, in patients with shoulder pain. *Int J Sports Phys Ther.* 2013;8:784–92.
- 35 Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain.* 2001;94:149–58.
- 36 Farrar JT, Portenoy RK, Berlin JA, Kinman JL, Strom BL. Defining the clinically important difference in pain outcome measures. *Pain.* 2000;88:287–94.
- 37 Moore A, Moore O, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics: use of pain intensity and visual analogue scales. *Pain.* 1997;69:311–5.
- 38 Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain.* 2004;8:283–91.
- 39 Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain.* 1983;17:45–56.
- 40 Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics. *Pain.* 1996;66:229–37.
- 41 Cook CE. Potential pitfalls of clinical prediction rules. *J Man Manip Ther.* 2008;16:69–71.