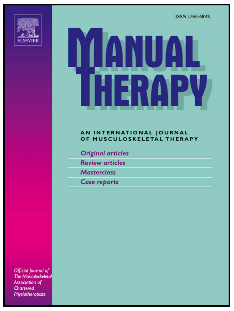
Accepted Manuscript



Criterion validity of manual assessment of spinal stiffness

Shane L. Koppenhaver, PhD, PT Jeffrey J. Hebert, PhD, DC, Senior Lecturer Greg N. Kawchuk, PhD, DC John D. Childs, PhD, PT, Director of Musculoskeletal Research Deydre S. Teyhen, PhD, PT Theodore Croy, PhD, PT Julie M. Fritz, PhD, PT

PII: S1356-689X(14)00119-2

DOI: 10.1016/j.math.2014.06.001

Reference: YMATH 1581

To appear in: *Manual Therapy*

Received Date: 13 November 2013

Revised Date: 27 May 2014

Accepted Date: 3 June 2014

Please cite this article as: Koppenhaver SL, Hebert JJ, Kawchuk GN, Childs JD, Teyhen DS, Croy T, Fritz JM, Criterion validity of manual assessment of spinal stiffness, *Manual Therapy* (2014), doi: 10.1016/j.math.2014.06.001.

Criterion validity of manual assessment of spinal stiffness

Shane L. Koppenhaver, PhD, PT[1]

Jeffrey J. Hebert, PhD, DC[3]

Greg N. Kawchuk, PhD, DC[4]

John D. Childs, PhD, PT[7]

Deydre S. Teyhen, PhD, PT[1]

Theodore Croy, PhD, PT[1]

Julie M. Fritz, PhD, PT[2]

Affiliations:

[1] Associate Professor, U.S. Army-Baylor University Doctoral Program in Physical Therapy, San Antonio, Texas.

[2] Associate Professor, University of Utah, College of Health. Clinical Outcomes Research Scientist, Intermountain Health Care, Salt Lake City, Utah.

[3] Senior Lecturer, School of Psychology and Exercise Science, Murdoch University, Perth, Australia.

[4] Associate Professor, Department of Physical Therapy, University of Alberta, Alberta, Canada.

[7] Director of Musculoskeletal Research, Department of Physical Therapy (MSGS/SGCUY), 81st Medical Group, Keesler Air Force Base, Biloxi, MS

This study was supported by the National Institutes of Health, National Center for Complementary and Alternative Medicine (R21 AT004221).

This study was approved by the Institutional Review Boards of the University of Utah.

The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Departments of the Army or Defence.

Corresponding Author:

Shane Koppenhaver

204 Tuttle Rd. San Antonio, TX, 78209

shanekoppenhaver@mac.com

210-722-3671

Abstract

Assessment of spinal stiffness is widely used by manual therapy practitioners as a part of clinical diagnosis and treatment selection. Although studies have commonly found poor reliability of such procedures, conflicting evidence suggests that assessment of spinal stiffness may help predict response to specific treatments. The current study evaluated the criterion validity of manual assessments of spinal stiffness by comparing them to indentation measurements in patients with low back pain (LBP). As part of a standard examination, an experienced clinician assessed passive accessory spinal stiffness of the L3 vertebrae using posterior to anterior (PA) force on the spinous process of L3 in 50 subjects (54% female, mean (SD) age = 33.0 (12.8) years, BMI = 27.0 (6.0) kg/m2) with LBP. A criterion measure of spinal stiffness was performed using mechanized indentation by a blinded second examiner. Results indicated that manual assessments were uncorrelated to criterion measures of stiffness (spearman rho = 0.06, p = 0.67). Similarly, sensitivity and specificity estimates of judgments of hypomobility were low (0.20-0.45) and likelihood ratios were generally not statistically significant. Sensitivity and specificity of judgments of hypermobility were not calculated due to limited prevalence. Additional analysis found that BMI explained 32% of the variance in the criterion measure of stiffness, yet failed to improve the relationship between assessments. Additional studies should investigate whether manual assessment of stiffness relates to other clinical and biomechanical constructs, such as symptom reproduction, angular rotation, quality of motion, or end feel.

Criterion validity of manual assessment of spinal stiffness

Introduction

Manual assessments of spinal stiffness have long been a cornerstone of the clinical examination for manual practitioners when assessing patients with spinal pain. Such assessments contribute to formulating a clinical diagnosis and often form the basis for treatment technique selection (Maitland 1986; Greenman 1996; Henderson 2012). For example, traditional manual therapy models use manual assessments of spinal stiffness to determine where to apply manual therapy, which technique to apply, as well as the direction and grade of application. A recent survey found that the great majority (98%) of manual physical therapists use manual assessments of spinal motion during their exam and base treatment decisions at least partially on their findings (Abbott et al. 2007). Additionally, emerging evidenced-based models of back pain management, such as the Treatment Based Classification System (Fritz et al. 2007; Hebert et al. 2011), use assessments of spinal stiffness to classify patients with low back pain (LBP) into clinically relevant subgroups.

Reliability of an examination procedure that is used for treatment decision-making is considered a prerequisite for its validity (Streiner and Norman 2003; Portney and Watkins 2008). The reliability of manual assessments of spinal stiffness has been extensively studied and systematically reviewed (Seffinger et al. 2004; van Trijffel et al. 2005; Stochkendahl et al. 2006; Schneider et al. 2008) Although estimates of reliability of manual assessment vary widely, with some studies reporting good reliability and others reports reliability no better than chance, systematic reviews report substantial qualitative deficits with the majority of these studies (Seffinger et al. 2004; van Trijffel et al. 2005; Stochkendahl et al. 2006). The latest systematic review focusing solely on inter-examiner reliability studies of intervertebral motion assessment of the lumbar and cervical spine (van Trijffel et al. 2005) found that only four out of 19 included studies were performed in patients with neck and back pain and that only three of the 19 studies included examiners that were blinded to each other's assessments. Although inconclusive due to these qualitative shortcomings, common findings of poor reliability, especially by higher quality studies (van Trijffel et al. 2005; Schneider et al. 2008) have led many researchers and clinicians to question the continued use of manual assessments of spinal stiffness as a part of the clinical examination (Wainner 2003; Seffinger et al. 2004; Landel et al. 2008).

Establishing validity for an examination procedure depends upon the procedure's intended use. Despite having poor reliability, some evidence suggests that manual assessment of spinal stiffness may have some predictive validity in determining which patients with back pain are likely to respond best to different treatments. Specifically the presence of stiffness among patients with LBP is predictive of clinical success after spinal manipulation (Flynn et al. 2002; Childs et al. 2004). Additionally patients with LBP judged as hypermobile have been found to do better with lumbar stabilization exercise program (Fritz et al. 2005b). These findings were the result of manual posterior to anterior assessments of spinal stiffness defined in the studies as at least one level (L1-L5) being rated as "hypomobile" or "hypermobile" on a 3-point scale (hypomobile, normal, hypermobile). Such findings suggest that manual assessments of spinal stiffness may be sufficiently valid to be useful components of the clinical examination (Wainner 2003).

Other studies that have investigated the validity of manual assessments of spinal stiffness have found less encouraging results. Several studies reported that choosing a manual therapy technique based on assessments of spinal stiffness results in no better outcomes than random selection (Chiradejnant et al. 2003a; Haas et al. 2003; Kanlayanaphotporn et al. 2009). Moreover, as a part of a population-based study, Leboeuf-Yde et al. (Leboeuf-Yde et al. 2002) found that manual assessments of spinal stiffness were not helpful in differentiating people with and without LBP. Although a "gold standard" measure of spinal stiffness is not well established, several studies have compared manual assessments of spinal motion to spinal motion assessed by imaging. Both Fritz (Fritz et al. 2005a) and Abbott (Abbott et al. 2005) found moderate agreement between manual assessments of spinal motion and motion during flexion and extension radiographs while Landel (Landel et al. 2008) found poor agreement between ratings of spinal motion between concurrent manual and MRI assessments.

A common limitation of the aforementioned criterion validity studies is that their criterions all measured only the amount of spinal motion, whereas clinicians assess both motion and resistance to motion, or spinal stiffness (Abbott et al. 2007; van Trijffel et al. 2009). Spinal indentation is a technique to quantify spinal stiffness using both force and linear displacement data. Previous studies comparing mechanized and manual assessments of spinal stiffness have only been performed in asymptomatic subjects and have generally found poor agreement unless examiners are specifically trained to match their assessments to the indentation results (Maher et al. 1998; Chiradejnant et al. 2003b). Therefore the primary purpose of this study was to evaluate the criterion validity of manual assessments of spinal stiffness by comparing such judgments to indentation measurements of spinal stiffness in patients with LBP. Additionally we explored the hypothesis that anthropometric characteristics of the patient (age, sex, and body mass index [BMI]) affect judgements made during manual assessments of spinal stiffness.

Materials and Methods

<u>Participants</u>

Volunteers with LBP were recruited from local physical therapy clinics and a university campus as a part of a larger study investigating the effects of spinal manipulation (Fritz et al. 2011). Participant inclusion and exclusion criteria are listed in Table 1 and were used to ensure a clinically relevant sample without contraindications to spinal manipulation. All participants reviewed and signed consent forms approved by the Institutional Review Board of the University and the rights of the participants were protected.

Procedures

After providing informed consent, participants completed several self-report measures and underwent a standard history and physical examination. The Numeric Pain Rating Scale (NPRS) was used to rate subjective back and leg pain intensity on a scale of 0 (no pain) to 10 (worst imaginable pain) (Childs et al. 2005). The modified Oswestry Disability Questionnaire (ODI) was used to quantify LBP-related disability (Fritz and Irrgang 2001). The standardized physical examination was similar to a typical clinical examination for LBP and included all of the tests and measures associated with the Treatment Based Classification System (Fritz et al. 2007; Hebert et al. 2011).

Index test

A licenced clinician with 8 years of clinical experience and who was blinded to the results of the indentation assessment performed the manual assessment of spinal stiffness. The spinous processes of L1-L5 were palpated on each prone participant. Each spinous process was marked to ensure consistent placement between manual assessment and the spinal indentation procedures. The examiner placed the region of the pisiform bone of his dominant hand on the posterior-most portion of the spinous process and then placed his non-dominant hand on top of the dominant hand for support. The participant was asked to relax as the examiner exerted a slow posterior to anterior (PA) force with both hands until he felt he reached the end of available spinal motion. The examiner then released approximately one half of his force and repeated several repetitions of the PA motion to assess the passive accessory spinal stiffness (see Figure 1). The stiffness of the each vertebral segment (L1-L5) was recorded as "hypomobile", "normal", or "hypermobile" based on based upon the clinician's perception of the amount of force used and the resultant segmental displacement. The presence or absence of pain was also recorded during the stiffness assessment of each level.

Criterion Standard

After the index test, spinal stiffness was quantified by an examiner blinded to the results of the manual stiffness assessment using a mechanized indentation device with established reliability (Stanton and Kawchuk 2009; Wong et al. 2013) and accuracy (Kawchuk et al. 2006). The indentation device consisted of a saddle tip attached to the terminal end of a linear stepping motor (Dual Motion Motor, HSI, Waterbury, CT) supported vertically by a rigid metal frame (Figure 2). Prior to undergoing indentation, participants were oriented to the machine and procedure including demonstration on a calibration device.

The transducer probe was positioned posterior to the L3 spinous process of each prone participant and slowly lowered until contacting the spinous process. L3 was chosen for the level of indentation on all participants as it is generally the segment that is most perpendicular with the indentation transducer. Initial pressure of the transducer was set at a comfortable level below 5 N which allowed normal respiration, but restricted participants from taking a full deep inhalation. The participant was then instructed to take a normal breath in and out and hold the breath at the end of exhalation. Towards the end of exhalation, the examiner started the indentation procedure at the preload of 5 N and progressed to a maximum load of 60 N before being automatically withdrawn. 60N was selected based on extensive pilot testing and was found to be an appropriate maximal load that adequately challenged the spine while remaining tolerable in our symptomatic sample. Linear indenter displacement was quantified by a rotary encoder and signals from the load cell (Measurement Specialties, Hampton, VA) and transformer were collected by customized LABview software (National Instruments, Austin, TX) at a collection rate of 200 Hz. Each indentation lasted approximately 5 second and was performed 3 times on each participant. If participants inhaled before the end of measurement, the repetition was repeated.

Force and linear displacement data were used to calculate spinal stiffness. Global Stiffness (GS) was the primary outcome and was calculated as the slope of the force/displacement curve between 5 and 60 N. Terminal Stiffness (TS) was additionally calculated as the instantaneous stiffness (N/mm) that occurred at the maximal indentation load. GS and TS measures were each averaged across the 3 indentation repetitions to reduce variability (Wong et al. 2013).

<u>Data Analyses</u>

All data were entered into and analyzed by IBM SPSS 21 (Chicago, IL). Descriptive statistics were performed on sociodemographic and health characteristics of the sample. The statistical significance and strength of relationship between manual assessments of spinal stiffness and indentation measures of spinal stiffness of L3 (GS and TS) were assessed using Spearman's rho correlation analysis.

Measurements were then dichotomized in order to calculate diagnostic utility estimates. Manual spinal stiffness outcomes were categorized into those judged by the clinician assessor to be "hypomobile" vs. "normal or hypermobile". Indentation measures (GS and TS) were dichotomized using two different distribution-based cut-offs of "stiffness", greater than vs. less than the sample mean and greater than vs. less than one standard deviation above the sample mean. Point estimates and 95% confidence intervals of sensitivity, specificity, and positive and negative likelihood ratios for each different criterion cut-off were calculated using an excel-based calculator downloaded from the Physiotherapy Evidence Database (www.pedro.org.au). Similar analyses could have been done for hypermobility by comparing those judged by the clinician assessor to be "hypermobile" vs. "normal or hypomobile". These analyses were not performed, however, as only 5 out of 50 participants were judged to be "hypermoble".

Lastly, we used stepwise hierarchical linear regression models to explore the hypothesis that anthropometric characteristics of the patient affect judgements made during manual assessments of spinal stiffness. The criterion measures of stiffness (GS and TS) served as the dependent variables. Age, sex, and BMI were entered into the model in the first step in a forward stepwise fashion. A significance value less than 0.05 was required for a variable to enter the model and greater than 0.10 to remove a variable from the model. Manual assessment of spinal stiffness was then force-entered into the second step.

Results

Fifty-one participants with LBP were recruited. Stiffness data were not captured on one post-partum participant due to the indenter exceeding its maximal displacement before reaching the terminal load of 60 N. Demographic and clinical characteristics for the remaining 50 participants are presented in Table 2.

Descriptive statistics for stiffness values alone and their correlation with manual judgements are presented in Table 3. Spearman's rho correlation coefficients were not statistically significant. Since the results for GS and TS were essentially the same for the other analyses, results are only presented for GS. As can be seen in Figure 3, participants who were judged to have "normal" intervertebral motion by manual assessments demonstrated the highest stiffness values.

Based on 2x2 contingency tables (Table 5), sensitivity, specificity, and positive and negative likelihood ratios of "hypomobility" are displayed in Table 6 using two cut-offs of stiffness

(GS). Regardless of the cut-off, sensitivity and specificity estimates were low (0.20-0.45) and likelihood ratios were generally not statistically different from 1, indicating a judgement of "hypomobile" does not significantly change the post-test likelihood of a participant being "stiff".

Of the anthropometric variables entered into the stepwise hierarchical linear regression, only BMI was retained after step one (Table 4) indicating that BMI was predictive of GS ($\beta = -0.566$, $p < 0.001$). Specifically BMI explained 32% of the variance associated with GS measures. After accounting for the relationship of BMI and GS, judgement of intervertebral stiffness made during manual assessment was not predictive of GS ($\beta = 0.006$, $p = 0.96$).

As an additional control measure to ensure that pain with manual assessment was not confounding other analyses, we examined the point-biserial correlation between "pain with L3 assessment" (painful vs. non-painful) and GS. Pain did not significantly correlate with GS ($r_{pb} = -0.17$, $p = 0.22$) suggesting that pain did not confound the relationship between manual assessment and indentation measures of stiffness.

Discussion

Clinicians who utilize manual therapy interventions frequently include judgments of spinal stiffness in their examination of patients with LBP (Abbott et al. 2007), presuming that increased stiffness indicates the need for a specific treatment (e.g. spinal manipulation). Re-assessment of spinal stiffness following treatment is then often used as a marker of having delivered a successful treatment if stiffness is perceived to have decreased (Tuttle 2009). The evolving paradigm of evidence-based practice dictates that clinicians focus on examination procedures that are both reliable and valid. Although manual assessments of spinal motion have most commonly been found to be unreliable (Seffinger et al. 2004; Stochkendahl et al. 2006; Schneider et al. 2008; van Trijffel et al. 2009) several studies suggest that such judgements are helpful with predicting benefit with specified treatments (Fritz et al. 2005b). To further explore the validity and diagnostic utility of manual assessments of spinal stiffness we compared such judgments to indentation measurements of spinal stiffness. Our results indicate that judgements of spinal hypomobility made during manual assessment are unrelated to, and are not helpful in identifying, alterations of spinal stiffness.

One possible explanation for our results is that manual assessments of spinal stiffness are inherently unreliable and inaccurate. Previous studies have shown that manual assessments show a great deal of variability in the magnitude and direction of applied force (Latimer et al. 1998; Caling and Lee 2001). This could explain the common research finding that such manual assessments are unreliable (Seffinger et al. 2004; Stochkendahl et al. 2006; Schneider et al. 2008; van Trijffel et al. 2009) and would support the notion that reliability is a prerequisite of validity. If manual assessments of spinal stiffness are simply unreliable and invalid, their continued use during clinical examination is difficult to justify. This conclusion has been reached by several other authors after reviewing the reliability literature (Troyanovich et al. 1998; Seffinger et al. 2004) and is consistent with studies that find a lack of association between assessments of spinal motion and clinical outcomes (Chiradejnant et al. 2003a; Haas et al. 2003; Kanlayanaphotporn et al. 2009).

Another possible explanation for our findings is that judgements based on manual assessments may be evaluating a different construct than spinal stiffness or may be evaluating multiple or combined constructs. In a recent survey of 466 U.S. and New Zealand manual physical therapists, Abbott et al. (2007) found that respondents reported assessing multiple constructs when performing manual assessments of spinal stiffness. “Pain response” was the construct reported most commonly as the most important, followed by “quality of resistance” (i.e. stiffness) and “quantity of translation” of the vertebrae. Many participants however, also reportedly evaluated “quality of end-feel” and “quality of motion path" during manual spinal assessments. Other studies comparing manual assessments of spinal motion to criterion measures assessed both angular spinal rotation and linear spinal displacement (Abbott et al. 2005; Fritz et al. 2005a). Because our criterion measure (indentation) measures only linear spinal stiffness, it could be that manual providers are detecting aspects of motion not included by the criterion test used in this study and/or making judgements on the relationship between constructs such as pain and stiffness.

It is also possible that manual providers consciously or unconsciously account for a patient's anthropometric characteristics (age, sex, and BMI) when manually assessing spinal stiffness, which may distort the relationship between manual assessments and indentation measures of spinal stiffness. It was anecdotally apparent during indentation measurements that larger individuals with more adipose tissue were measured as substantially "less stiff". Therefore, we explored the hypothesis that age, sex, and BMI may affect the relationship between manual assessment of spinal stiffness and indentation measures of spinal stiffness. We found that BMI, but not age or sex, was related to indentation measures of spinal stiffness. Moreover, we found that manual judgements of spinal stiffness did not relate to indentation measures after accounting for the relationship between BMI and indentation measures.

Perhaps the most salient limitation of the current study is that only one aspect of manual assessment of the spine was evaluated. The examiner simply rated passive accessory vertebral motion as "hypomobile", "normal", or "hypomobile" and did not attempt to qualify different components of spinal motion such as quality of motion, resistance to motion, or end feel. Although this is the same methodology of assessing spinal motion used in studies that have found predictive validity in determining which patients with back pain are likely to respond best to different treatments (Flynn et al. 2002; Childs et al. 2004; Fritz et al. 2005b), it is possible that having providers specifically focus on the force-displacement curve on which the mechanized spinal stiffness assessments are based would have resulted in better agreement between manual assessments and indentation measures. Additionally, since judgements of hypermobility were relatively infrequent (5 out of 50 participants) we limited our sensitivity, specificity, and likelihood ratio analyses to hypomobility rather than performing them on both hypomobility and hypermobility. Although this limits us from making quantitative conclusions about diagnostic utility of judgements of hypermobility, the graph of assessments of each participant (Figure 3), suggest that manual judgements of spinal stiffness are poor discriminators of criterion stiffness, regardless of the category breakdown.

Another limitation of the current study is that, regardless of the specific location of the subject's back pain, spinal stiffness was always measured at the L3 vertebrae. Although we are unaware of such evidence, it is possible that the assessments of spinal stiffness would better relate to criterion stiffness measures if measured at the most focal areas of pain. Finally, the application parameters of both the manual assessment and the indentation assessment were developed to optimize each separately rather than be standardized together. To maximize generalizability, the examiner performed the manual assessment in an identical fashion that he had previously used in his 8 years of clinical practice. Similarly, the parameters of the indentation measures were selected to most accurately measure the force-displacement curve at a tolerable level of force in participants with LBP. There may have been small differences in several parameters between the manual assessment and indentation measures including the amount of load, rate of loading, and padding between the 2 tables that may have adversely affected their relationship. While some of these parameters may have adversely affected the agreement between the index and criterion tests, the differences would likely be systematic and would not affect the ordinal relationship (correlation) between the two measures and would also be more representative of clinical practice.

Future research should further investigate the clinical utility of manual assessment of spinal stiffness. If additional studies verify the predictive validity of manual assessments, future research should investigate whether manual assessment of spinal stiffness relates to other constructs of spinal motion such as quality of motion or end feel and explore alternative methods of objectively quantifying these different constructs.

Bibliography

Abbott J, Flynn T, Fritz J, Hing W, Reid D, Whitman J. Manual physical assessment of spinal segmental motion: Intent and validity. Man Ther [Internet]. 2007; Available from: file://localhost/Users/shanekoppenhaveMBP/Documents/Bookends/Attachments/Abbott%20et%20al%202007.pdf

Abbott JH, McCane B, Herbison P, Moginie G, Chapple C, Hogarty T. Lumbar segmental instability: a criterion-related validity study of manual therapy assessment. BMC musculoskeletal disorders. 2005;6(1):56.

Caling B, Lee M. Effect of direction of applied mobilization force on the posteroanterior response in the lumbar spine. Journal of Manipulative and Physiological Therapeutics. 2001 Feb;24(2):71–8.

Childs JD, Fritz JM, Flynn TW, Irrgang JJ, Johnson KK, Majkowski GR, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. Ann Intern Med. 2004;141(12):920–8.

Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. Spine. 2005;30:1331–5.

Chiradejnant A, Maher C, Latimer J, Stepkovitch N. Efficacy of “therapist-selected” versus “ randomly selected” mobilisation techniques for the treatment of low back pain: a randomised controlled trial. Aust J Physiother. 2003 a;49(4):233–41.

Chiradejnant A, Maher CG, Latimer J. Objective manual assessment of lumbar posteroanterior stiffness is now possible. J Manipulative Physiol Ther. 2003 b Jan;26(1):34–9.

Flynn T, Fritz J, Whitman J, Wainner R, Magel J, Rendeiro D, et al. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. Spine. 2002;27(24):2835–43.

Fritz J, Koppenhaver S, Kawchuk G, Teyhen D, Hebert J, Childs J. Preliminary Investigation of the Mechanisms Underlying the Effects of Manipulation: Exploration of a Multivariate Model Including Spinal Stiffness, Multifidus Recruitment, and Clinical Findings. Spine (Phila Pa 1976). 2011;36(21):1772–81.

Fritz JM, Cleland JA, Childs JD. Subgrouping patients with low back pain: evolution of a classification approach to physical therapy. J Orthop Sports Phys Ther. 2007;37(6):290–302.

Fritz JM, Irrgang JJ. A Comparison of a Modified Oswestry Disability Questionnaire and the Quebec Back Pain Disability Scale. Phys Ther. 2001;81(2):776–88.

Fritz JM, Piva SR, Childs JD. Accuracy of the clinical examination to predict radiographic instability of the lumbar spine. Eur Spine J. 2005 a;14:743–50.

Fritz JM, Whitman JM, Childs JD. Lumbar spine segmental mobility assessment: an examination of validity for determining intervention strategies in patients with low back pain. Arch Phys Med Rehabil. 2005 b;86(9):1745–52.

Greenman PE. Principles of Manual Medicine. Baltimore, Maryland: Williams & Wilkins; 1996.

Haas M, Groupp E, Panzer D, Partna L, Lumsden S, Aickin M. Efficacy of cervical endplay assessment as an indicator for spinal manipulation. Spine. 2003;28(11):1091–6; discussion 1096.

Hebert JJ, Koppenhaver SL, Walker BF. Subgrouping Patients With Low Back Pain. Sports Health: A Multidisciplinary Approach. 2011;3(6):534–42.

Henderson CNR. The basis for spinal manipulation: Chiropractic perspective of indications and theory. Journal of Electromyography and Kinesiology [Internet]. 2012 Apr [cited 2012 Apr 17]; Available from: http://linkinghub.elsevier.com/retrieve/pii/S1050641112000582

Kanlayanaphotporn R, Chiradejnant A, Vachalathiti R. The immediate effects of mobilization technique on pain and range of motion in patients presenting with unilateral neck pain: a randomized controlled trial. Arch Phys Med Rehabil. 2009;90(2):187–92.

Kawchuk GN, Liddle TR, Fauvel O, Johnston C. The accuracy of ultrasonic indentation in detecting simulated bone displacement: a comparison of three techniques. Journal of manipulative and physiological therapeutics. 2006;29(2):126–33.

Landel R, Kulig K, Fredericson M, Li B, Powers C. Intertester reliability and validity of motion assessments during lumbar spine accessory motion testing. Phys Ther. 2008;88(1):43–9.

Latimer J, Lee M, Adams RD. The effects of high and low loading forces on measured values of lumbar stiffness. J Manipulative Physiol Ther. 1998 Apr;21(3):157–63.

Leboeuf-Yde C, van Dijk J, Franz C, Hustad SA, Olsen D, Pihl T, et al. Motion palpation findings and self-reported low back pain in a population-based study sample. J Manipulative Physiol Ther. 2002 Feb;25(2):80–7.

Maher CG, Latimer J, Adams R. An investigation of the reliability and validity of posteroanterior spinal stiffness judgments made using a reference-based protocol. Physical Therapy. 1998;78(8):829–37.

Maitland GD. Vertebral Manipulation. Oxford, England: Butterworth Heinemann; 1986.

Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice (3rd Edition). Upper Saddle River, NJ: Pearson Prentice Hall; 2008.

Schneider M, Erhard R, Brach J, Tellin W, Imbarlina F, Delitto A. Spinal palpation for lumbar segmental mobility and pain provocation: an interexaminer reliability study. J Manipulative Physiol Ther. 2008;31(6):465–73.

Seffinger MA, Najm WI, Mishra SI, Adams A, Dickerson VM, Murphy LS, et al. Reliability of spinal palpation for diagnosis of back and neck pain: a systematic review of the literature. Spine. 2004;29(19):E413.

Stanton TR, Kawchuk GN. Reliability of assisted indentation in measuring lumbar spinal stiffness. Man Ther. 2009 Apr;14(2):197–205.

Stochkendahl MJ, Christensen HW, Hartvigsen J, Vach W, Haas M, Hestbaek L, et al. Manual Examination of the Spine: A Systematic Critical Literature Review of Reproducibility. Journal of Manipulative and Physiological Therapeutics. 2006 Jul;29(6):475–485.e10.

Streiner D, Norman G. Health Measurement Scales: A Practical Guide to Their Development and Use. New York, NY: Oxford University Press; 2003.

Van Trijffel E, Anderegg Q, Bossuyt PMM, Lucas C. Inter-examiner reliability of passive assessment of intervertebral motion in the cervical and lumbar spine: A systematic review. Manual Therapy. 2005 Nov;10(4):256–69.

Van Trijffel E, Oostendorp RAB, Lindeboom R, Bossuyt PMM, Lucas C. Perceptions and use of passive intervertebral motion assessment of the spine: a survey among physiotherapists specializing in manual therapy. Man Ther. 2009 Jun;14(3):243–51.

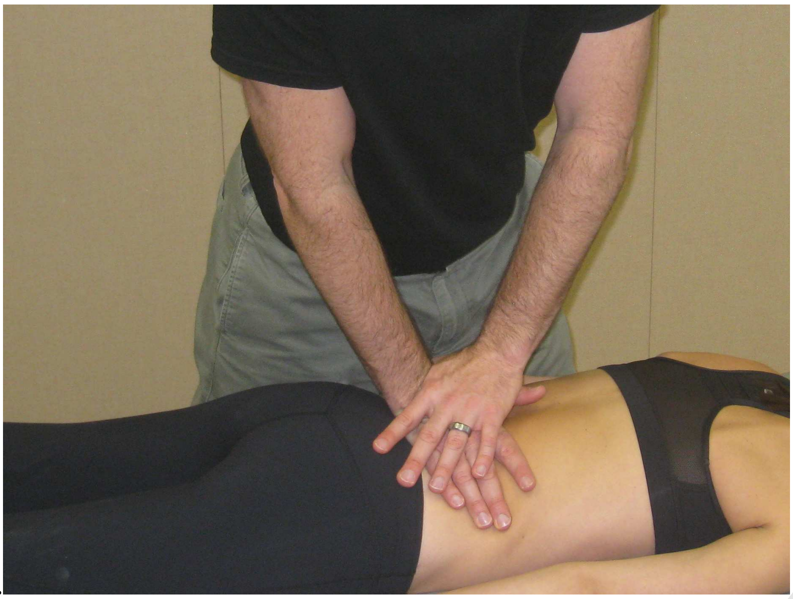
Troyanovich SJ, Harrison DD, Harrison DE. Motion palpation: it's time to accept the evidence. J Manipulative Physiol Ther. 1998 Oct;21(8):568–71.

Tuttle N. Is It Reasonable to Use an Individual Patient's Progress After Treatment as a Guide to Ongoing Clinical Reasoning? Journal of Manipulative and Physiological Therapeutics. 2009 Jun;32(5):396–403.

Wainner RS. Reliability of the clinical examination: how close is "close enough"? J Orthop Sports Phys Ther. 2003 Sep;33(9):488–91.

Wong AYL, Kawchuk G, Parent E, Prasad N. Within- and between-day reliability of spinal stiffness measurements obtained using a computer controlled mechanical indenter in individuals with and without low back pain. Manual Therapy [Internet]. 2013 Mar [cited 2013 Mar 14]; Available from: http://linkinghub.elsevier.com/retrieve/pii/S1356689X13000325

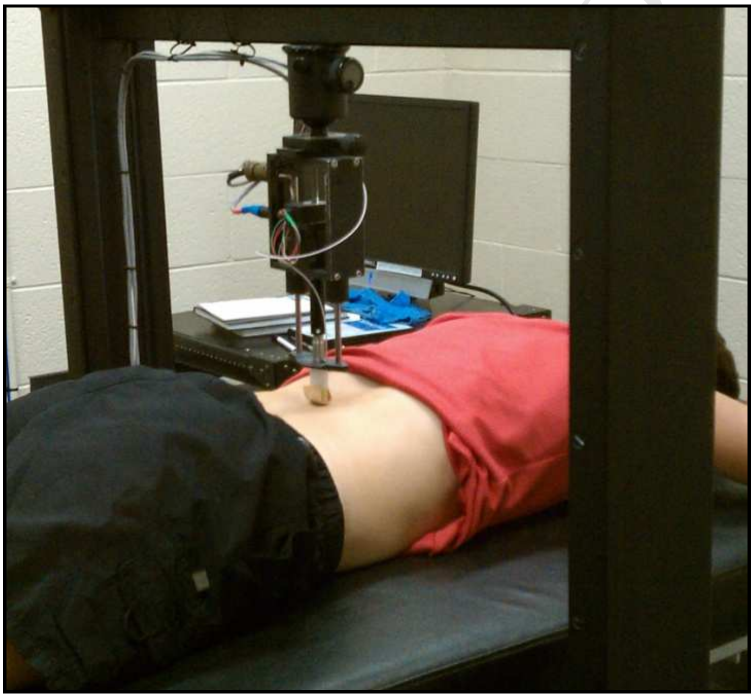
1 **FIGURE 1.** Posterior to anterior mobilization used as index test of manual assessment of spinal stiffness



2 .

3

4 **FIGURE 2.** Mechanized indentation device used as the criterion standard measure of spinal stiffness



5

6 **TABLE 1.** Study inclusion and exclusion criteria

Inclusion Criteria	Exclusion Criteria
Back pain located between the 12th rib and buttocks, that in the opinion of the screening examiner, was originating from the lumbar region	Neurogenic pain defined by either a positive ipsilateral or contralateral straight leg raise (reproduction of symptoms at ≤ 45°) or reflex, sensation, or strength deficits in a pattern consistent with nerve root compression
Between the age of 18 and 60 years	Osteoporosis
	Prior surgery to the lumbosacral spine
Modified Oswestry Disability score at least 20%	Medical 'red flags' of a potentially serious condition including cauda equina syndrome, major or rapidly progressing neurological deficit, fracture, cancer, infection, or systemic disease
Ability to lie prone and supine for a minimum of 20 minutes	Prior spinal manipulation to the lumbosacral spine or trunk muscle stabilization exercises performed in the previous 4 weeks

TABLE 2. Demographic and Clinical Characteristics (n=50).

Characteristic	
Age	33.0 (12.8) years
Sex	52.0% female
BMI	27.0 (6.0) kg/m2
Numeric Pain Rating*	4.9 (1.6)
Oswestry Disability Score	32.2 (12.0) %
Prior History of LBP	88.0% yes
Duration of Symptoms	184 (41, 758)† days
Distribution of Symptoms	26.0% with leg pain

Numbers represent mean (standard deviation) unless otherwise indicated
*Reports the average of the worst, best and current scores for pain over the last 24 hours
† Median (interquartile range).
BMI: Body Mass Index, LBP: low back pain

TABLE 3. Descriptive statistics of spinal stiffness values stratified by manual judgement of spinal stiffness

Manual PA Judgement	Global Stiffness	Terminal Stiffness
	Mean (SD)	*Mean (SD)*
Hypomobile (n = 24)	5.19 (1.28)	4.06 (1.58)
Normal (n = 21)	6.18 (1.84)	4.89 (2.04)
Hypermobile (n = 5)	4.63 (1.12)	3.49 (1.15)
Spearman's Rho Correlation (p-value)	0.06 (0.67)	0.07 (0.63)

FIGURE 3. Global Stiffness (GS) measures of each individual categorized by judged intervertebral stiffness. Interpolation line represents mean of each category.

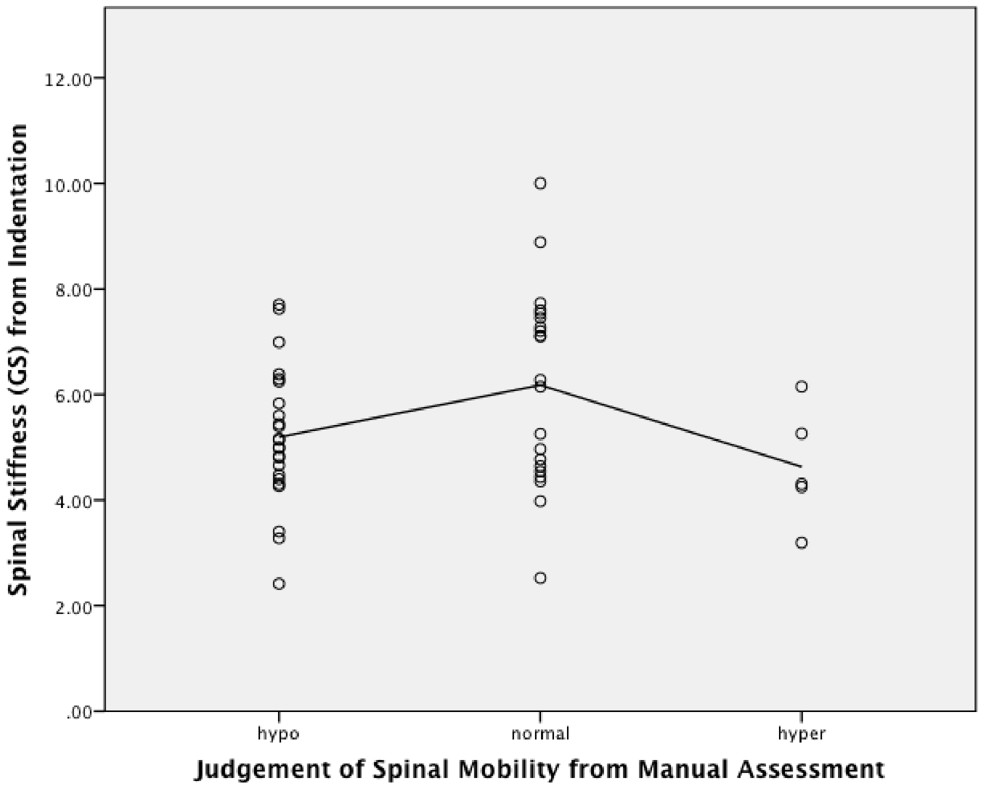


TABLE 4. Hierarchical linear regression analysis predicting criterion measure of spinal stiffness (GS)

Variables	Standardized β coefficient	Significance of β coefficient	Adjusted R2 Change
Body Mass Index (kg/m2)	-0.566	< 0.001	0.321
Manual Assessment of Spinal Stiffness	0.006	0.958	< 0.001

TABLE 5. 2x2 contingency tables for the two reference standards used to evaluate the manual assessment of stiffness.

	Spinal Indentation		
	> Mean stiffness	≤ Mean stiffness	Total
Manual assessment			
Rated hypomoble	8	16	24
Rated normal or hypermobile	13	13	26
Total	21	29	50

	Spinal Indentation		
	> +1SD stiffness	≤ +1SD stiffness	Total
Manual assessment			
Rated hypomoble	2	22	24
Rated normal or hypermobile	8	18	26
Total	10	40	50

TABLE 6. Diagnostic accuracy of manual assessment of spinal stiffness to detect spinal stiffness (GS)

Criterion Standard	Sensitivity (95% CI)	Specificity (95% CI)	LR + (95% CI)	LR- (95% CI)
> Mean stiffness	0.38 (0.21, 0.59)	0.45 (0.28, 0.62)	0.69 (0.37, 1.31)	1.38 (0.82, 2.33)
> +1SD stiffness	0.20 (0.06, 0.51)	0.45 (0.31, 0.60)	0.36 (0.10, 1.30)	1.78 (1.12, 2.82)

GS: Global stiffness