

Spinal palpatory diagnostic procedures utilized by practitioners of spinal manipulation: annotated bibliography of content validity and reliability studies

Michael Seffinger, DO*
Alan Adams, DC**
Wadie Najm, MD—
Vibian Dickerson, MD—
Shira I Mishra, MD, PhD §
Sib, lle Reinsch, PhD §§
Linda Murph, , MLS §§

... 30 ...
... I ...
2001.
(JCCA 2003; 47(2):93-109)

... 30 ...
... I ... 2001.
(JACC 2003; 47(2):93-109)

KEY WORDS : manual therap, , palpation, spinal manipulation.

MOTS CL S : traitement manuel, palpation, manipulation spinale.

Introduction

Professions that employ manual manipulation procedures use their own terminology to describe the diagnostic entity that responds to manipulation therapy.¹ Spinal neuromusculoskeletal dysfunction is the term used in our paper to encompass these various terms employed by the different disciplines. Spinal neuromusculoskeletal dysfunction refers to an alteration of spinal joint position, motion characteristics and related palpable paraspinal soft tissue changes. Spinal palpation diagnostic procedures typically entail static palpation of anatomical landmarks for symmetry, palpation of spinal vertebral joints before, during and after active and passive motion tests, and spinal and paraspinal soft tissue palpation assessment for abnormalities or altered sensitivity.

Outcomes and effectiveness of manipulation treatments rely, partly, on the validity and reliability of the palpation procedures used to diagnose spinal neuromusculoskeletal dysfunction. Investigation of the validity and reliability of spinal palpation diagnostic tests has been in progress for the past 70 years. A complete review and analysis of these studies is lacking. A preliminary evaluation reveals an inconsistency in the focus, methodology, palpation procedures and statistical analysis used. Focused narrative reviews from this literature have been previously published.²⁻¹⁰

This paper is an annotated bibliography of primary research studies on the content validity and reliability of

M6 Beal MC, Goodridge JP, Johnston WL, McConnell DG. Inter-examiner agreement on long-term patient improvement: an exercise in research design. J Am Osteopath Assoc 1982; 81:322–328.

The study evaluated long-term inter-examiner reliability on diagnosing aggravation or improvement of musculoskeletal conditions of the spine. Three osteopathic physician faculty, specializing in manipulation performed unspecified spinal palpation, evaluation and manual treatment procedures on 3 patients over 17 months. Palpation, evaluation included regional and segmental motion testing and palpation of paraspinal soft tissue. Line graphs using a plus-minus scale demonstrated overall agreement of findings on 2 out of 3 patients. The authors also stated that an improvement in palpation test results correlated with patients' reported symptomatic improvement.

M7 Johnston W, Hill J, Elkiss M, Marino R. Identification of stable somatic findings in hypertensive subjects by trained examiners using palpation examination. J Am Osteopath Assoc 1982; 81:830–836.

The study evaluated the inter-examiner reliability of spinal

inter-examiner reliability, the scores were 0.19 to 0.17. There was poor inter-examiner reliability for all segments with fair to moderate intra-examiner reliability at the L1-2 and the L4-5 segments.

M18 Leboeuf C, Gardner V, Carter A, Scott T. Chiropractic examination procedures: A reliability and consistency study. J Austral Chiropractor Assoc 1989; 19:101-104.

The study investigated intra- and inter-examiner reliability for certain chiropractic tests: pain on spinous process palpation, interspinous ligament palpation and spinous process percussion, and motion palpation, sign of the rising thumb and resilience on extension in the lumbar spine. Two chiropractic students examined 39 subjects with chronic low back pain. The ability of examiners to agree on the presence/absence of positive findings in these

M28 Tuchin P, Hart C, Johnson C, et al. Inter-examiner reliability of chiropractic evaluation for cervical spine problems – a pilot study. Part 1: Graduates from one institution. Australian Chiropractic & Osteopathy 1996; 5:23–29.

The study evaluated the reliability of examiners in palpation of the cervical spine. Eight experienced chiropractors, using individual palpator methods including static and/or motion palpation, and vertebral springing, examined 53 asymptomatic volunteer student subjects for cervical spine dysfunction. There was poor inter-examiner agreement with C6 being the segment of the highest disagreement.

M29 Phillips DR, Twomey LT. A comparison of manual diagnosis with a diagnosis established by a uni-level lumbar spinal block procedure. This study was presented in part at the 8th Biennial Conference of the MPAA in 1993. Manual Therapy 1996; 1:82–87.

This study investigated the inter-examiner reliability and validity of lumbar spine manual palpation in the diagnosis of patients with low back pain using a randomized crossover design with a prospective and retrospective part. Two manipulative physiotherapists evaluated 63 asymptomatic and 9 asymptomatic volunteer subjects for abnormal quantity and quality of passive intervertebral motion and vertebral response to digital pressure. The authors used percent agreement, Kappa and weighted Kappa analysis to determine inter-examiner reliability. There was poor inter-therapist reliability for motion ratings (weighted Kappa ranged from .015 to 0.32) and vertebral response to pressure (Kappa ranged from .016 to 0.28).

0.308 for inter-examiner and intra-examiner reliability, respectively. The authors concluded that there was poor inter and intra-examiner reliability, for the craniovertebral side bending passive motion test.

M33 Maher CG, Latimer J, Adams R. An investigation of the reliability and validity of posteroanterior spinal stiffness judgments made using a reference-based protocol. *Phys Ther* 1998; 78:829–837.

The reliability portion of this study assessed the agreement in stiffness estimates in the lumbar spine. Two blinded experienced physical therapists examined 40 asymptomatic volunteers at L3, using their preferred palpation method. Interclass correlation coefficient (ICC) with 95% confidence intervals was 0.5–0.62 with a low standard error of measurement.

M34 Hawk C, Phongphua C, Bleecker J, Swank L, Lopez D, Rubley T. Preliminary study of the reliability of assessment procedures for indications for chiropractic adjustments of the lumbar spine. *J Manipulative Physiol Ther* 1999; 22:382–389.

The study evaluated inter- and intra-examiner reliability in assessing indications for chiropractic adjustment of the lumbar spine. Four licensed chiropractic examiners (2 with 20+ and 2 with 3 or less years of experience), examined 18 (2 symptomatic and 16 asymptomatic) volunteer subjects. Examiners were trained in a standardized flexion-distraction technique. They also used manual assessment procedures used in everyday clinical practice including 1) hypermobility of each segment, 2) changes in tissue texture or tension of the skin and underlying tissue, 3) palpable temperature changes, and 4) tenderness elicited on palpation. Intra-examiner reliability, Kappa scores ranged from 0.17 to 0.85. For intra-examiner reliability, there was considerable variation by segment and among the four examiners. Inter-examiner reliability scores ranged from 0.42 to 0.44. The authors

procedures. Overall, intra-examiner agreement on decision to manipulate ranged from 73 to 92 percent, with Kappa Values of 0.13 to 0.73. Inter-examiner reliability, as low at lower thoracic and lumbar spinal levels, with the mean percent agreement ranging from 48 to 83% and Kappa Values ranging from 0.16 to 0.27.

examiner reliability, between the two trained examiners was greater (93% agreement; Kappa = 0.85) than for either of these two with the (third) untrained examiner, who used his own test procedures (82% and 84% agreements; Kappa = 0.57 and 0.61 respectively). Intra-examiner reliability, for one of the trained examiners was good (91% agreement; Kappa = 0.78).

Pain or sensitivity provocation procedures

P1 Waddell G, Main CJ, Morris EW, et al. Normality and reliability in the clinical assessment of backache. *BMJ (Clin Res Ed)* 1982; 284:1519–1523.

The study evaluated the inter-examiner reliability of history, and physical examination procedures and clinical assessment in patients with back pain. Five orthopedic surgeons examined 810 patients with backache. This included elicitation of lumbar tenderness by spinal palpation. Most study groups compared two examiners. The unweighted Kappa scores for reliability of spinal palpation for tenderness on 8 patients was 1.0 ($p < 0.001$). The unweighted kappa scores for reliability on physical examination ranged from 0.41 to 1.0 and on psychosocial and behavioral assessment from 0.27 to 0.94 ($p < 0.05$).

P2 DeBoer K, Harmon R, Tuttle C, Wallace H. Reliability study of detection of somatic dysfunctions in the cervical spine. *J Manipulative Physiol Ther* 1985; 8:9–16.

See M10

P3 Viikari-Juntura E. Inter-examiner reliability of observations in physical examinations of the neck. *Phys Ther* 1987; 1526–1532.

The study assessed inter-examiner reliability of palpation procedures of the cervical spine. A physical medicine and rehabilitation specialist (physiatrist) and a physical therapist examined 69 consecutive symptomatic patients using a conventional neurological evaluation, palpation, and provocative tests for pain, numbness and paresthesias. Agreement on palpation for tenderness as reported for 51 subjects using empirical value of Kappa scores and proportion of significant agreement (p_s). The Kappa score was fair for upper spinous processes at 0.47 (p_s 0.56), as well as for lower spinous processes at 0.52 (p_s 0.67); however, as poor for right cervical paraspinal soft tissues at 0.24

(p_s 0.33). For left cervical paraspinal soft tissues the prevalence was less than 10% so Kappa was not used (p_s 0.00).

P4 Bolin P, Keating J, Brist J, Denver G. Inter-examiner reliability of palpatory evaluations of the lumbar spine. *Am J Chiropractic Med* 1988; 1:5–11. See M14

P5 Leboeuf C, Gardner V, Carter A, Scott T. Chiropractic examination procedures: A reliability and consistency study. *J Austral Chiropractor Assoc* 1989; 19:101–104. See M18

P6M Keating JC, Jr. Bergmann TF, Jacobs GE, Finer BA, Larson K. Inter-examiner reliability of eight evaluative dimensions of lumbar segmental abnormality. *J Manipulative Physiol Ther* 1990; 413:463–70. See M19

P7M Nice DA, Riddle DL, Lamb RL, Mayhew TP, Rucker K. Intertester reliability of judgments of the presence of trigger points in patients with low back pain. *Arch Phys Med Rehabil* 1992; 73:893–898.

The study evaluated inter-examiner reliability in assessing the presence of trigger points in the lumbar spine region. Ten experienced physical therapists evaluated 50 patients with low back pain using the Traubell and Simon assessment examination using pain as the endpoint. The Kappa score for inter-examiner reliability ranged from 0.29 to 0.38; percent agreement ranged from 76% to 79%; the observed proportion of positive agreement ranged from 0.43 to 0.52. The authors concluded that there was poor inter-examiner reliability in the assessment of the presence of trigger points in patients with low back pain.

P8 Bolin PD, Haas M, Meyer JJ, Kassak K, Nelson C, Keating JC Jr. Inter-examiner reliability of eight evaluative dimensions of lumbar segmental abnormality: Part II. *J Manipulative Physiol Ther* 1993; 16:363–374.

The study evaluated inter-examiner reliability using several measurements including lumbar spinal palpation procedures. Three experienced chiropractors examined 28 symptomatic patients with chronic low back pain. Palpa-

tion for osseous pain produced percent agreement that ranged from 79% to 96% with Kappa coefficients ranging from 0.48 to 0.98. Palpation for soft tissue pain produced percent agreement ranging from 75% to 93% with Kappa coefficients ranging from 0.40 to 0.79. Good reliability, as demonstrated for inter-examiner provocative palpation, procedures for elicitation of both osseous and paraspinal soft tissue pain in the study population.

P9 Richter T and Lawall J. Reliability of diagnostic findings in manual medicine. *Manuelle Medizin*, 1993; 31:1–11.

See M22

P10 Hubka MJ, Phelan SP. Inter-examiner reliability of palpation for cervical spine tenderness. *J Manipulative Physiol Ther* 1994; 17:591–595.

The study evaluated the inter-examiner reliability of palpation for cervical spine tenderness using a within subjects (repeated measures) design. Two experienced chiropractors examined 30 patients with mechanical neck pain. Inter-examiner reliability, as assessed by percent agreement was 76.6% with a Kappa score of 0.68. The authors found that manual palpation of the cervical spine for tenderness is a reliable examination tool.

P11 Maher C, Adams R. Reliability of pain and stiffness assessments in clinical manual lumbar spine examination. *Phys Ther* 1994; 74:801–811.

**P19 Schöps P, Siebert U, Schmitz U, Friedle AM,
Beyer A. Reliabilität nichtinvasiver diagnostischer
Untersuchungsmethoden zur Erfassung**

cient (ICC) at the 95% confidence interval as calculated.
Intra-examiner reliability was poor to moderate for both

tion. Two experienced chiropractors examined 42 asymptomatic chiropractic students in the sitting and prone positions. The examiners used a skin-marking pen to identify the spinous processes. Intra-examiner percent agreement at L1 with the subject sitting was 55% for one examiner and 39% for the other. At L4, intra-examiner percent agreement for both examiners was 62%. Inter-examiner percent agreement was better at L4 (sitting 79%; prone 81%) than at L1 (sitting 55%; prone 69%) with the subjects in either the sitting or prone positions.

L4 Binkley J, Stratford PW, Gill C. Inter-rater reliability of lumbar accessory motion mobility testing. Phys Ther 1995; 75:786–792.

See M24

L5 McKenzie AM, Taylor NF. Can Physiotherapists locate lumbar spinal levels by palpation? Physiother 1997; 83:235–239.

The study evaluated intra-examiner and inter-examiner reliability in locating lumbar spinal levels by palpation. Three physiotherapists (intra-examiner) and 14 physiotherapists (inter-examiner) examined 10 volunteer subjects, using their preferred method of palpation. Kappa scores for intra-examiner reliability were 0.61 to 0.90. Kappa scores for inter-examiner reliability was 0.28. There was good to excellent intra-examiner reliability, but poor inter-examiner reliability when palpating for lumbar spine levels.

L6 Downey BJ, Taylor NF, Niere KR. Manipulative physiotherapists can reliably palpate nominated lumbar spinal levels. Man Ther 1999; 4:151–156.

The study assessed inter-examiner reliability in palpating lumbar spine levels. Three pairs of experienced physical therapists palpated 60 patients with low back pain, marking the mid-point of a randomly nominated spinous process. Almost perfect overall agreement was achieved among all three pairs in locating the nominated level (weighted Kappa = 0.92).

Unspecified types of tests

U1 McConnell DG, Beal MC, Dinnar U, et al. Low agreement of findings in neuromusculoskeletal examinations by a group of osteopathic physicians using their own procedures. J Am Osteopath Assoc 1980; 79:441–450.

The study examined the inter-examiner reliability in neuromuscular examination procedures, including spinal palpation. Six osteopathic physicians specializing in manipulation using their (unspecified) customary palpation procedures examined 21 asymptomatic volunteers. Results revealed low inter-examiner reliability on segmental location and intensity of findings. The authors inferred that inter-examiner agreement would likely improve if the examiners first agreed upon the following: a) the areas to be examined; b) the test procedures to be used; c) the method of quantifying the intensity of the findings; and d) the method of recording.

U2 Beal M, Dvorak J. Palpatory examination of the spine: a comparison of the results of two methods and their relationship to visceral disease. Man Med 1984; 1:25–32.

The study evaluated inter-examiner agreement using two methods of spinal palpation: the conventional American osteopathic and the Manual Medicine Society of Scotland methods. Two physicians specializing in spinal man-

ipulation examined 60 patients of routine -

id on the condition segment, dysfunction.

tion of painful upper cervical joint dysfunction. The findings of the chief investigator of the trial were compared with those of each of the other 6 experienced independent physiotherapists. The therapists examined 40 symptomatic (headache and neck pain) and asymptomatic volunteer subjects using their own personal test procedures. Additionally, some of the independent examiners were tested against each other. There was complete agreement (Kappa = 1.0) in six pairs of examiners and excellent agreement (Kappa = 0.78 and Kappa = 0.8) between two pairs. Percent agreement was 70% for inter-examiner reliability on the most dysfunctional joint in symptomatic patients.