

# Inter-Examiner Reliability of Motion Palpation of the Lumbar Spine:

## A REVIEW OF QUANTITATIVE LITERATURE

Editorial Staff

This study was authored by Joseph C. Keating, Jr., Ph.D., and was first published in Vol. 2, No. 3 -- September 1989 issue of the American Journal of Chiropractic Medicine. It is reprinted with the permission of both Dr. Keating and the American Journal of Chiropractic Medicine.

**Abstract:** In the past decade at least seven quantitative studies of the inter-examiner reliability of motion palpation of the lumbar spine have been reported in the refereed literature. In this survey several characteristics of these seven papers were noted, including type of motion palpation, qualifications and experience of examiners, characteristics of subjects, levels of the lumbar spine which were examined, units of measurement, types of inferential statistic, and the appropriateness of conclusions. On the basis of this review, no strong claims for the objectivity of lumbar motion palpation are justified at this time. Although several measurement strategies deserve replication, the current literature generally demonstrates marginal-to-no reliability, limited numbers of examiners, and overreliance on asymptomatic students as palpatory subjects. (Am J Chiropractic Med 1989;3:107-110)

**Key words:** chiropractic, inter-examiner reliability, literature review.

### Introduction

Palpatory methods of evaluating segmental motion in the lumbar spine are widely used by practitioners of spinal adjusting, manipulation and mobilization. The identification of altered motion (e.g., fixations, hypermobility) is believed to provide important information for the choice of site and method for manual interventions. Among chiropractors, altered motion findings are thought to vary with pathologic changes<sup>1,2</sup> and to indicate probable malposition and subluxation.<sup>2</sup>

In spite of the widespread use of motion palpatory methods, relatively little critical attention has been paid to the measurement characteristics of these procedures. Despite the acknowledged "face validity"<sup>3</sup> of motion palpation, these methods have rarely been used to quantitatively monitor spinal changes within clinical trials. Little information is available to support the contention that palpably detectable alterations in lumbar segmental motion are associated with other suspected or detectable components of joint dysfunction or subluxation (e.g., pain, muscle tension, nerve interference). Similarly, the significance (or lack thereof) of motion palpation findings in determining the "manipulable lesion" has not been critically, quantitatively evaluated. On the other hand, some data are available to support the stability (i.e., test-retest or intra-examiner reliability) of some palpatory measurements of altered motion.<sup>4</sup>

This paper provides a review of an additional measurement characteristic of lumbar palpation: objectivity. The objectivity or inter-examiner reliability of a measurement system provides an indication of how well two or more observers who evaluate the same subject at the same time can agree with one another. A measurement system which demonstrates strong inter-examiner

reliability may not yield valid observations, since multiple examiners may err in similar fashion and thereby agree. However, inter-examiner studies play an important screening function in clinical research, since a measurement system which fails to demonstrate objectivity is unlikely to demonstrate validity. Inter-examiner studies of motion palpation of the lumbar spine provide a template upon which future subluxation-detection studies and clinical trials of the adjustment can and should be planned.

At this time at least seven papers on the inter-examiner reliability of lumbar motion palpation have been published (or accepted for publication) in the refereed scientific literature (see Table 1). These reports have varied in terms of palpatory methods, examiners and qualifications, sample characteristics, spinal levels examined, units of analysis, methods of inferential statistical analysis, and results obtained.

Three types of lumbar motion palpation have been studied. Gonella et al<sup>4</sup> employed a form of "active motion palpation"<sup>1</sup> wherein the examiner guides the subject through an active range of motion while digitally probing the spinous processes for movement. Love and Brodeur<sup>7</sup> employed a scanning procedure in which the back of the examiner's hand (proximal phalanges) was used to stress spinal areas from L4/L5 through T1/T2, and thereby to identify the stiffest or "most hypomobile" region. This procedure provides a screen for subsequent, more specific segmental evaluations of motion. Five papers have reported on the objectivity of passive motion palpation.<sup>5,6,8,9,10</sup> "Passive motion palpation"<sup>1</sup> typically involved digital challenges to the spinous processes with the subject positioned so that the joint under examination was at the end of its normal range of motion. Examiners then noted a restricted or hypermobile range of joint motion and/or a hand end-feel.<sup>5,8,10</sup>

Examiners have included doctors of chiropractic (DCs), physical therapists (PTs), DC students, and PT students. Experience with motion palpation procedures varied from a few semesters to many years. Some investigators reported pre-study, unblinded rehearsals of standardized protocols by participating clinicians. Four of the seven studies employed only two examiners.

Relatively healthy, young asymptomatic chiropractic students have been the most frequent subjects of inter-examiner reliability studies of lumbar motion palpation, but PT students and low back pain patients have also been studied. However, of the 303 subjects studied in these seven reports, only 44 definite low back patients (15 percent) were included. Sample sizes have ranged from 5 to 100 subjects.

In all but one investigation<sup>7</sup> each lumbar segmental level was examined and a rating made for each. Love and Brodeur's<sup>7</sup> procedure identified the stiffest of 16 thoracolumbar joints rather than rating each motion unit. All of the studies were examined, segments L1/L2 through L4/L5, and most also considered the motion of the thoracolumbar and lumbosacral junctions.

In four studies<sup>6,8,9,10</sup> the clinicians were asked to make a dichotomous choice upon palpating each segment -- that a fixation was present or absent at that particular joint. On the other hand, Gonella et al.<sup>4</sup> employed a 13-point scale (0 through 6 in half-unit intervals, where 0 = ankylosed, 3 = normal, and 6 = unstable to rate each segment. Similarly, Jull and Bullock<sup>5</sup> employed a 5-point rating system (1 = slightly hypermobile, 2 = normal, 3 = slightly, 4 = moderately, and 5 = markedly hypomobile) to judge each lumbar joint at each of its end-of-ranges. As suggested above, Love and Brodeur's<sup>7</sup> paradigm involved selection of a spinal area (segment) rather than rating each

joint. Their ordinal unit of measurement was the position along the spine where the greatest hypomobility was detected.

Inferential statistical analyses have included linear correlational methods (Pearson's "r"; index of association) and concordance coefficients (Kappa). These procedures<sup>11</sup> provide a means of evaluating whether the observed agreement or predictability between examiners' findings could be written off to chance, or must be considered so unusual that the null hypothesis (i.e., no real concordance or correlation between clinicians) is rejected. In the clinical sciences, the concordance or linear predictability between doctors must be so clear-cut it would be expected, by chance only, 1 time out of 20 or less frequently ( $p < .05$ ).

Love and Brodeur<sup>7</sup> applied correlational methods to their ordinal data, but could find no significant linear relationship between examiners. Jull and Bullock<sup>5</sup> employed Pearson's method to test the linear correlation between doctors' mobility ratings, and reported  $r = .82$  to  $r = .94$  across several directions of movement (flexion, extension, left or right rotation, left or right lateral flexion, posteroanterior glide). Unfortunately, these authors did not report their coefficients in terms of individual segmental levels, and did not report the probability associated with each correlation coefficient. Neither is it clear what each coefficient means nor how it was computed. Relatedly, their impressive report of 86 percent exact agreement and not more than a one-grade discrepancy between doctors in the remaining 14 percent of paired observations, seems to support their endorsement of the inter-examiner reliability of lumbar motion palpation. However, Jull and Bullock<sup>5</sup> did not report the distributions of paired findings for the 10 pain-free subjects in their sample. Since their scale was limited to five grades, an error limit of plus or minus one grade is probably too liberal. Suppose that one examiner rated almost every joint "two" (normal), and the second examiner agreed 86 percent of the time that the joint was normal. In a pain-free population, this level of agreement on a five-point scale might well represent less than chance concordance (an unweighted Kappa for five-choice concordance could be applied). Summarily, Jull and Bullock's<sup>5</sup> analysis does not support their positive conclusions.

Three investigations have employed the Kappa coefficient<sup>12</sup> to evaluate the significance of agreement between doctors' dichotomous ratings (fixed, not fixed) of lumbar segmental motion.

Boline et al<sup>9</sup> reported that two examiners were able to agree to a weak but statistically significant degree when palpating at the thoraco-lumbar junction (T12/L1) and the mid-lumbar spine (L3/L4), but not at other lumbar segmental levels. Keating et al<sup>10</sup> could not find consistently significant, pair-wise concordance among three experienced examiners who employed passive motion palpation, although "marginal reproducibility"<sup>12</sup> was noted at L4/L5 and L5/S<sup>1</sup> for two out of three examiner-pairs. Mootz et al<sup>6</sup> could find no significant agreement between examiners for the presence/absence of fixations at L1/L2 through L5/S.

Two of the papers reviewed here did not use inferential statistics. Gonella et al<sup>4</sup> apparently employed the "eye-ball" method to determine that lumbar palpators could not agree on segmental mobility ratings. Bergstrom and Curtis<sup>6</sup> relied upon descriptive statistics (percent-agreement) to evaluate concordance, and inappropriately concluded that their dichotomous rating system was reliable.<sup>3</sup> In fact, their results cannot be interpreted based on the data they report.

## Conclusion

Taken together, these papers do not provide strong evidence of the inter-examiner reliability of

lumbar motion palpation. However, Jull and Bullock's<sup>5</sup> impressive demonstrations of reliability at all segmental levels and in all directions of motion deserves to be replicated in larger, symptomatic samples. Similarly, the marginal findings noted by Boline et al<sup>9</sup> and Keating et al<sup>10</sup> justify further study in larger patient groups.

These conclusions stand at odds with the beliefs of many in the profession. Dishman<sup>2</sup>, for example, suggests that "Inter-examiner reliability studies indicate that a standard method of motion palpation is quite feasible and accurate." Williams<sup>13</sup> considers inter-examiner reliability research to be a "misdirection of efforts," and seems to believe that many if not most diagnostic methods have demonstrated satisfactory reproducibility across examiners. These uncritical views stand as barriers to more widespread evaluation of motion palpation and other clinical assessment procedures, for they undercut the motivation to investigate by assuring doctors that we already know what works in chiropractic.

"Motion palpation" may or may not include valid, reliable and useful methods of evaluating lumbar segmental dysfunction. Based upon currently available data, however, these procedures have not received the unambiguous research support that would justify strong claims. Clinicians should include a degree of skepticism in their use of these methods.

#### Acknowledgement

I wish to thank Dr. Gert Bronfort for his assistance in retrieving materials for this paper.

#### References

1. Cassidy, J.D., Potter, G.E. "Motion Examination of the Lumbar Spine." J Manipulative Physiol Ther 1979; 2(3):151-8.
2. Dishman, R.W. "Static and Dynamic Components of the Chiropractic Subluxation Complex." J Manipulative Physiol Ther 1988; 11(2): 98-107.
3. Keating, J.C.; Boline, P.D. "Letter to the Editor." J Manipulative Physiol Ther 1988; 11(5):443-4.
4. Gonella, C.; Paris, S.V.; Kutner, M. "Reliability in Evaluating Passive Intervertebral Motion." Phys Ther 1982; 62(4):436-44.
5. Jull, G., Bullock, M. "A Motion Profile of the Lumbar Spine in An Aging Population Assessed By Manual Examination." Physiotherapy Practice 1987; 3:70-81.
6. Mootz, R.D.; Keating, J.C.; Kontz, H.P.; Milus, T.B.; Jacobs, G.E. "Inter- and Intra-Observer Reliability of Passive Motion Palpation of the Lumbar Spine." J Manipulative Physiol Ther, accepted for publication 6:88.
7. Love, R.M.; Brodeur, R.R. "Inter- and Intra-Examiner Reliability of Motion Palpation for the Thorarolumbar Spine." J Manipulative Physiol Ther 1987; 10(1):1-4.
8. Bergstrom, E.; Courtis G. "An Inter- and Intra-Examiner Reliability Study of Motion Palpation in Lateral Flexion in the Seated Position." Euro J Chiropractic 1986; 34:121-41.
9. Boline, P.D.; Keating, J.C.; Brist, J.; Denver, G. "Inter-Examiner Reliability of Palpatory Evaluations of the Lumbar Spine." Am J Chiropractic Med 1988; 1(1):5-11.
10. Keating, J.C.; Bergmann, T.F.; Jacobs, G.E.; Finer, B.; Larson, K. "The Objectivity of a Multi-dimensional Index of Lumbar Segmental Abnormality." J Manipulative Physiol Ther. Accepted for publication 11:88.
11. Keating, J.C. "Several Strategies for Evaluating the Objectivity of Measurements in Clinical Research and Practice." J Can Chiropractic Assoc Sept. 1988; 32(3):133-8.
12. Rosner, B. "Fundamentals of Biostatistics." 2nd ed., Boston: PWS Publishers, 1986.
13. Williams, S.E. "Research and Responsibility." Chiropractic Research Jrl 1988; 1(2):1-2.

\*Associated Professor, Palmer College of Chiropractic-West, Sunnyvale, California. Paper presented at the Fourth Annual Conference on Research and Education of the California Chiropractic Foundation, June 29-July 2, 1989, Monterey, California and the International Chiropractors' Association's Scientific Symposium on Spinal Biomechanics, May 19-21, 1989, St. Louis, Missouri.

Paper submitted April 10, 1989.

AUGUST 1990