Research report

Reliability of physical examination to assess asymmetry of anatomical landmarks indicative of pelvic somatic dysfunction in subjects with and without low back pain

Adrian Kmita, Nicholas P. Lucas*

School of Biomedical and Health Sciences, University of Western Sydney, Narellan Road, Campbelltown NSW 1797, Australia

Abstract

Objective: To investigate the intra-examiner and inter-examiner reliability of physical examination to identify asymmetry of selected anatomical landmarks indicative of pelvic somatic dysfunction in subjects with and without low back pain using experienced osteopaths and final year students of osteopathy.

Methods: Four examiners (two students, two osteopaths) examined a sample of symptomatic ($n = 5$) and asymptomatic ($n = 4$) subjects for symmetry of anatomical landmarks indicative of pelvic somatic dysfunction. Two assessments of symmetry and alignment of the posterior superior iliac spine (PSIS), sacral sulcus, sacral inferior lateral angle (ILA) in posterior−anterior (ILA-P) and inferior−superior (ILA-I) directions, anterior superior iliac spine (ASIS), and medial malleoli were performed on every subject by all four examiners. Intra-examiner and inter-examiner reliability was analysed with kappa (κ) and reported in conjunction with observed agreement (P_{o}).

Results: Estimates of intra-examiner reliability ranged from $\kappa = -0.29$ to 1.0 (PSIS $\kappa = -0.29$ to 0.39; sacral sulcus $\kappa = -0.28$ to 0.83; ILA-P $\kappa = -0.29$ to 0.44; ILA-I $\kappa = -0.29$ to 0.34; ASIS $\kappa = 0.25-0.63$; medial malleoli $\kappa = 0.20-1.0$) and were higher than estimates of inter-examiner reliability. Inter-examiner reliability estimates ranged from $\kappa = -0.38$ to 0.51 (PSIS $\kappa = -0.38$ to 0.35; sacral sulcus $\kappa = -0.34$ to 0.26; ILA-P $\kappa = -0.18$ to 0.51; ILA-I $\kappa = -0.13$ to 0.36; ASIS $\kappa = -0.13$ to 0.50; medial malleolus $\kappa = -0.05$ to 0.49). The median observed agreement between examiners for each anatomical landmark ranged from 33 to 50%. Osteopaths were more reliable on measures of the inferior lateral angle (ILA-P), while students were more reliable on measures of the sacral sulcus.

Conclusion: In this study, the reliability of physical examination for anatomical landmarks indicative of pelvic somatic dysfunction was generally found to be low. Differences between the reliability of experienced osteopaths and final year osteopathy students were negligible. Examiners were most reliable in their assessment of the ASIS and medial malleolus; however, these estimates were not consistent and were too low to be considered clinically useful.
© 2008 Published by Elsevier Ltd.

Keywords: Osteopathic medicine; Diagnosis; Physical examination; Somatic dysfunction

1. Introduction

In the practice of osteopathic medicine, palpatory diagnosis is emphasised for the identification of somatic

dysfunction in the spine and pelvis.[1] Somatic dysfunction is believed to involve a functional disturbance of the musculoskeletal system[2] which may lead to the production of symptoms or frank pathology.[1,3−5] Osteopaths in clinical practice are obligated to ensure that they use the most accurate diagnostic methods available, including those used to identify somatic dysfunction. In order to determine the usefulness of a test, two important

* Corresponding author. Tel.: +61 4620 3100.

E-mail addresses: adriankmita@yahoo.com (A. Kmita), n.p.lucas@uws.edu.au (N.P. Lucas).

diagnostic characteristics are evaluated: reliability and validity. Reliability is a measure of the extent of agreement between two and more test outcomes.[6] Test validity is a measure of how well the test performs in comparison to a reference standard.[7] Tests that lack sufficient reliability are not useful, as they do not provide a consistent measure of the variable of interest (e.g. pelvic landmark symmetry).[6]

The issue of diagnostic reliability in osteopathic medicine is vexatious, as the majority of studies demonstrate that commonly used physical examination tests used to identify somatic dysfunction are not reliable, or have only poor or slight reliability. Seffinger et al.[8] conducted a systematic review of the literature to assess the inter-examiner and intra-examiner reliability of spinal palpatory diagnostic procedures. The 12 articles rated most highly in this review revealed that landmark palpation, pain provocation and motion tests have acceptable reliability, however, were not always reproducible by other examiners under similar conditions. Overall, factors such as the examiners' discipline, experience level, consensus on procedure used, training just before the study, or use of symptomatic subjects, did not appear to improve reliability. Stochkendahl et al.[9] conducted a systematic literature review on the reliability of spinal palpation and concluded that palpation for pain is reproducible at a clinically acceptable level among observers. However, the reproducibility of static palpation, motion palpation and soft tissue changes was not clinically acceptable.

It may be argued that somatic dysfunction is dynamic and may change with assessment, and that this provides a reason to explain why studies consistently demonstrate unsatisfactory reliability.[8,10−15] However, this is a tenuous argument for proponents of these tests to make, primarily because it is inconsistent with the concept of somatic dysfunction. An aetiological time frame is central to the concept of somatic dysfunction in health and disease. Only a dysfunction that is consistently present can be considered as potentially aetiological in the production of symptoms. Many biological variables fluctuate over time, and are dynamic rather than static. Only when a reliable method of measuring the variable is developed, can normative data be established from which to determine abnormal deviations. If somatic dysfunction is unstable, inconsistently present, or modifiable with simple assessment techniques, how can it be found, how can it be fixed, and how can its correction be determined? It is reasonable to acknowledge that the body is dynamic, and that no two physical examinations will be identical in the same patient. However, this is true of all physical interactions and is exactly why a test must be reliable in order to be useful.

The skill and experience of examiners may also be a potential reason for the poor performance of these tests. Students may not be taught a systematic approach to physical examination; or insufficient time and practice may be the reason for poor reliability. Practitioners may diverge in their approach and employ different palpatory techniques over time. These reasons do not remove the requirement that tests should be reliable, even though they are legitimate reasons for why test reliability might be poor.

The primary aim of anatomical landmark assessment is to determine asymmetry between bilateral structures and deviation of normal alignment of unpaired structures. The identification of pelvic somatic dysfunction is described in most textbooks of osteopathy, and is defined in the Glossary of Osteopathic Terminology under *sacral, pubic, and innominate somatic dysfunction*.[2] Fryer et al.[11] investigated the reliability of anatomical landmark assessment for the identification of pelvic somatic dysfunction in asymptomatic volunteers, using trained and untrained student examiners. Examiners were found to have low reliability, with trained examiners obtaining the highest results for the medial malleolus ($\kappa = 0.31$) and ASIS ($\kappa = 0.24$), followed by the PSIS ($\kappa = 0.08$) and ILA ($\kappa = 0.04$). To further the work of Fryer et al.,[11] it is necessary to investigate the reliability of these tests as used by experienced osteopaths in symptomatic subjects. In addition, statistical estimates of reliability are improved if there is a reasonably equal spread of positive and negative cases represented in the sample.[16−19] This occurs because inadequate or over-representation of the sign of interest (e.g. asymmetry) in the sample can influence kappa paradoxically toward reporting low values in the presence of high examiner agreement. Since there is no reference standard test for the identification of pelvic somatic dysfunction, there is no objective means by which investigators can recruit a sample in which 50% have signs of somatic dysfunction, and 50% do not. In order to improve the likelihood of an even spread of positive and negative findings, investigators may include both symptomatic and asymptomatic subjects in their sample. While the *symptom* of pain is not the *sign* that these tests are designed to identify, it is consistent with osteopathic theory that those with symptoms are more likely to have signs of somatic dysfunction in comparison to those without symptoms. In order to further establish the reliability of anatomical landmark assessment for signs of pelvic somatic dysfunction, we conducted a reliability study with experienced osteopaths and final year osteopathy students in a mixed cohort of subjects with and without low back pain.

2. Methods

2.1. Subjects

Subjects were recruited from the outpatient teaching clinic of the University of Western Sydney using advertisements and word of mouth invitation. Potential subjects