Journal of Bodywork & Movement Therapies (2012) 16, 289–293

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/jbmt

ASSESSMENT RELIABILITY STUDY

# The inter-rater and intra-rater reliability of passive physiological accessory movement assessment of lumbar spine in novice manual therapists

**Mangesh Deore, M.Sc.** [a,b,*], **Stephen May, PhD** [b]

[a] *Manchester NHS Clinical Assessment and Treatment Services, United Kingdom*
[b] *Sheffield Hallam University, Collegiate Crescent Campus, 36 Collegiate Crescent, Sheffield S10 2BP, United Kingdom*

Received 31 October 2011; received in revised form 10 December 2011; accepted 20 December 2011

**KEYWORDS**
Accessory movements;
Lumbar spine;
Reliability;
Novices

**Summary** Passive Physiological Accessory Movements (PPAVMs) are commonly used assessment and treatment techniques in patients with low back pain. Many physiotherapists, including novices, consider PPAVMs an important tool for assessment and treatment of low back pain. Reliability is important as a judgement on the reproducibility of assessment procedures between therapists. However, the reliability of PPAVMs seems to have some problems, and reliability of PPAVMs has not yet been established amongst novice manual therapists. This study aimed at investigating inter-rater and intra-rater reliability of PPAVMs in novice physiotherapists. Fifty two healthy participants were recruited for the study. PPAVMs were applied by two novice physiotherapists and accessory movements were assessed, and both the raters were blinded from each others' findings to avoid bias. The mobility was graded on a three point scale with grade 1 being considered as hypomobile, 2 as normal and 3 as hypermobile. This procedure was performed on all five lumbar segments. Each participant was assessed on the same day for inter-rater reliability, and for intra-rater reliability the participant was assessed by one rater a week later. Kappa ($\kappa$) was calculated for all the levels of lumbar spine which ranged between 0.01 and 0.30 for inter-rater reliability and 0.24 to 0.57 for intra-rater reliability. Percentage exact agreement was also computed which showed a range between 38.4% and 57.6%. The values of '$\kappa$' showed poor intra-rater and inter-rater reliability. However, further research is advisable in order to assess the role of experience in reproducibility of PPAVMs.


* Corresponding author. 3rd Floor, 111 Piccadilly, Manchester M1 2HY, United Kingdom. Tel.: +44 (0) 1143484270, +44 (0) 7944010057.
*E-mail address:* Mangesh.Deore@careuk.com (M. Deore).

1360-8592/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jbmt.2011.12.006

## Introduction

Low back pain (LBP) is one of the commonest musculoskeletal disorders. Evidence suggests that the LBP shows point prevalence of around 30% in the adult population (Anderson, 1999). LBP often has social and economical implications and leads to significant absenteeism from work (Paatelma et al., 2005). Treatment plans should vary according to the assessment findings and therefore it is important to conduct detailed examination of the patients with LBP. Assessment for LPB usually includes history-taking and physical examination, which should provide important information to shape management (Magee, 2002). The physical examination can consist of palpatory methods of assessment of which Passive Physiological Accessory Movements (PPAVMs) are one of the tools for the judgement of stiffness or mobility between two vertebrae (Magee, 2002; Maitland et al., 2005). PPAVMs are delivered by positioning the patient in prone and a postero-anterior (PA) pressure is applied by the therapist by using a pisiform contact on the spinous process under consideration (Maitland et al., 2005). Intervertebral mobility may guide formulation of the treatment plan for LBP (Fritz et al., 2005).

It is important to ensure that procedures that determine treatment are reliable. An assessment technique needs to be reliable to be considered as valid (Sim and Wright, 2005). Reliability is a measure of reproducibility of findings achieved by the same or different assessors. Consistent readings are important as any variations observed may confuse the treatment plan. Therefore, it is important to establish the reliability of PPAVMs as these techniques are commonly used by manual therapists and other clinicians.

There have been previous studies which attempted to assess the reliability of accessory as well as inter-segmental mobility of the lumbar spine (Maher and Adams, 1994; Binkley et al., 1995; Trijffel et al., 2005; Landel et al., 2008). Maher and Adams (1994) used a 11-point scale ranging from 0 to 10 in ascending order of mobility (hypomobility to hypermobility) to measure the mobility between lumbar vertebrae. The results showed poor reliability of intra-rater and inter-rater findings. The Intraclass Correlation Coefficient (ICC) values ranged from 0.03 to 0.37 for all the lumbar segments. These values seem to be below the acceptable statistical levels of reproducibility. However, the scale used by this study seems to be too extensive as clinically the mobility is categorised broadly into hypomobility, normal and hypermobility (Maitland et al., 2005). Poor reliability results can be attributed to the extensive scale of grading mobility as there seemed to be minimal differences between two adjacent grades.

A similar study was conducted by Binkley et al. (1995) who used a 9 point ascending scale ranging from 1 to 9. 'Excess severe motion' and 'no motion' were considered at the extremes of this scale. The raters were experienced manual physiotherapists with minimum of seven years of experience. The results showed poor inter-rater reliability with ICC of 0.25. Kappa analysis showed low levels of agreement. Again the scale was very extensive, which is likely to lead to more disagreements. Moreover, the long measurement scales used by Maher and Adams (1994) and Binkley et al. (1995) are seldom used in clinical practice.

More recent studies (Landel et al., 2008) showed conflicting evidence. Landel et al. (2008) conducted a reliability study to identify the most and least mobile segments in the lumbar spine. Two raters who assessed the mobility of lumbar segments had a minimum of 15 years of experience in manual therapy. It was found that intra-rater reliability for the least mobile segments was good (kappa = 0.71), whereas it was poor (kappa = 0.29) for identifying the most mobile segments. The raters noted only the most and least mobile segments and no other mobility rating scale was used.

Trijffel et al. (2005) conducted a systematic review to investigate the reliability of PPAVMs, and recommended that more high quality studies were needed to establish evidence of reliability. In another systematic review May et al. (2006) also found that majority of assessment techniques for the lumbar spine showed poor inter-rater reliability, and most studies were of low or moderate quality.

Previous studies recruited experienced physiotherapists/manual therapists for the assessment of reliability of PPAVMs; however, these studies did not evaluate reliability in novice manual therapists. There is a need to assess the psychometric properties of PPAVMs in novice manual therapists as some clinical environments may lack availability of experienced therapists. Moreover, PPAVMs are skill dependent techniques and experience may play some role in assessing the intervertebral mobility.

This study considered the impact of experience on the reliability of techniques as many novice manual therapists practise PPAVMs widely in clinical decision making and treatment of LBP. It can be expected that many manual therapists may use these techniques despite the dearth of evidence around the role of experience on reliability of PPAVMs.

Methodological quality also needs to be considered strongly in order to gain real insight regarding the reliability of PPAVMs. Therefore, this study aimed at establishing the intra-rater and inter-rater reliability of PPAVMs in novice manual therapists.

## Methods

### Research design

A test-retest design was used to establish the reliability of PPAVMs. The repeated findings by one rater were considered for intra-rater reliability, whereas findings by two independent raters were considered for inter-rater reliability. The study was approved by Sheffield Hallam University Health and Social Care Ethics Committee.

### Participants

The participants included for this study were healthy volunteers between the ages of 18—40 years from Sheffield Hallam University, UK, who responded to invitations to participate via e-mail, direct contact or telephone communication. The voluntary participants were excluded if they were experiencing any pain in the low back area, had sustained any recent trauma or fractures, or were