Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Manual Therapy 11 (2006) 28–39
MANUAL THERAPY
www.elsevier.com/locate/math

Original article

The inter-examiner reliability of a classification method for non-specific chronic low back pain patients with motor control impairment

W. Dankaerts[a,b,*], P.B. O'Sullivan[a], L.M. Straker[a], A.F. Burnett[a], J.S. Skouen[c,d]

[a]School of Physiotherapy, Curtin University, Bentley 6102, WA, Australia
[b]Department of Rehabilitation Sciences and Physiotherapy, Ghent University, Ghent, Belgium
[c]The Outpatient Spine Clinic, Haukeland University Hospital, Bergen, Norway
[d]Section of Physiotherapy Science, Department of Public Health and Primary Health Care, Faculty of Medicine, University of Bergen, Norway

Received 15 June 2004; received in revised form 23 December 2004; accepted 16 February 2005

Abstract

The importance of classifying chronic low back pain (LBP) patients into homogeneous sub-groups has recently been emphasized. This paper reports on two studies examining clinicians ability to agree independently on patients' chronic LBP classification, using a novel classification system (CS) proposed by O'Sullivan. In the first study, a sub-group of 35 patients with non-specific chronic LBP were independently classified by two 'expert' clinicians. Almost perfect agreement (kappa-coefficient 0.96; %-of-agreement 97%) was demonstrated. In the second study, 13 clinicians from Australia and Norway were given 25 cases (patients' subjective information and videotaped functional tests) to classify. Kappa-coefficients (mean 0.61, range 0.47–0.80) and %-of-agreement (mean 70%, range 60–84%) indicated substantial reliability. Increased familiarity with the CS improved reliability. These studies demonstrate the reliability of this multi-dimensional mechanism-based CS and provide essential evidence in a multi-step validation process. A fully validated CS will have significant research and clinical application.

Keywords: Agreement; Classification; Chronic low back pain; Motor control impairment; Reliability

1. Introduction

Low back pain (LBP) is one of the most common and costly musculo-skeletal pain syndromes, affecting up to 80% of people at some point during their lifetime (Katz, 2002; van Tulder et al., 2002; Ehrlich, 2003; Woolf and Pfleger, 2003). The re-occurrence rate for LBP is high and these disorders often develop into a chronic fluctuating problem with intermittent flares (Croft et al., 1998; Burton et al., 2004). It has been stated that caring for chronic LBP (CLBP), is one of the most difficult and unrewarding problems in clinical medicine (Leclere et al., 1990) as no approach to diagnosis or treatment has been shown to be clearly definitive or effective.

One explanation offered for the inability to identify effective treatments is the lack of success in defining sub-groups of patients who are most likely to respond to a specific treatment approach (Leboeuf-Yde et al., 1997; Bouter et al., 1998). Indeed, it has been proposed that the 'LBP-group' conceals a large heterogeneous group of patients (McKenzie, 1981; Delitto et al., 1995; Spitzer et al., 1995; Borkan et al., 1998; Bouter et al., 1998; O'Sullivan, 2000; Leboeuf-Yde & Manniche, 2001). Any specific treatment applied to a falsely assumed homogenous sample may result in either failure to respond or

*Corresponding author. School of Physiotherapy Building 408, Curtin University of Technology, GPO Box U1987, Perth, WA 6845, Australia. Tel.: +61 08 9266 3667; fax: +61 08 9266 3699.
E-mail address: W.Dankaerts@exchange.curtin.edu.au (W. Dankaerts).

1356-689X/$ - see front matter
doi:10.1016/j.math.2005.02.001

aggravation of the disorder (Binkley et al., 1993; Fritz et al., 2000; Leboeuf-Yde & Manniche, 2001; Fritz et al., 2003).

The perceived need to accurately classify LBP into homogenous sub-groups to facilitate treatment to be tailored for specific disorders, led to an international forum on LBP ranking the accurate and reproducible characterization of sub-groups of patients with LBP as the top research priority (Borkan et al., 1998).

In general, criteria to classify can be defined as belonging to specific theoretical constructs or dimensions of the domain being classified (Bailey 1994; Ford et al., 2003). The shift from thinking about LBP as a patho-anatomical disorder, to viewing LBP as a multi-factorial bio-psycho-social disorder is now well accepted (Borkan et al., 2002). As a consequence of this, the different dimensions relevant to classifying the domain of LBP are patho-antomical, signs and symptoms, psychological and social (Waddell, 1987; Ford et al., 2003). For LBP, several classification systems (CSs) from a multitude of perspectives have been proposed. Recent systematic reviews highlight that the multi-dimensional nature of LBP is not reflected in most CS (Ford et al., 2003; McCarthy et al., 2004).

The authors propose that for non-specific CLBP, there is a special need for a mechanism-based CS acknowledging the bio-psycho-social dimensions of this disorder (Woolf et al., 1998; Ford et al., 2003; O'Sullivan, 2004a). When the mechanism or cause of a disorder is known, as long as it is amenable for treatment, treatment of the cause is usually considered more effective than treating its individual signs and symptoms (Zimny, 2004).

Table 1 gives an overview of the more commonly used categories and dimensions used to classify LBP patients and their limitations. It is increasingly clear that unidimensional CS's have limited clinical utility as do not adequately reflect the nature of LBP nor lead to its effective management. For CLBP there is no validated mechanism-based multi-dimensional CS. The development and testing of new CSs based on a multi-dimensional construct has been recommended (Riddle, 1998; Borkan et al., 2002; O'Sullivan, 2004a).

Recently, O'Sullivan (2000, 2004b) proposed a novel CS based on multiple dimensions for a sub-group of patients with NS-CLBP and clinical signs of motor control impairment (MCI). There is indeed considerable evidence documenting the presence of MCI in subjects with NS-CLBP, although, the nature of the impairment is highly variable (Hodges & Richardson, 1997; O'Sullivan et al., 1997; Hodges and Richardson, 1998; O'Sullivan et al., 1998; Hodges and Richardson, 1999; O'Sullivan, 2000; Sahrmann, 2001) and many mechanisms have been postulated for how pain may alter motor planning (Biedermann et al., 1991; Luoto et al., 1999; Hodges, 2001). O'Sullivan's (2000, 2004b) CS has been described in detail elsewhere but in brief it proposes (based on very strict inclusion and exclusion criteria) that a sub-group of patients with NS-CLBP exists (Table 2). These patients have impairments in the control of their lumbar spine that expose them to repeated stress and strain, thereby providing a basis for ongoing pain. Five distinct clinical patterns were proposed (Appendix A) based on a specific direction of MCI and the hypothesized mechanism underlying the pain disorder (O'Sullivan 2000, 2004b).

Whilst O'Sullivan's CS appears conceptually coherent, its reliability and validity should be established before its widespread use in clinical practice and research. The validation of a CS is a multi-step process (Woolf et al., 1998; Ford et al., 2003; Fritz et al., 2003; Dankaerts et al., 2004; O'Sullivan, 2004a) within which establishing the inter-examiner reliability is a crucial step. Therefore, the aim of the studies reported in this paper was to determine the inter-examiner reliability of this clinical method of classification for NS-CLBP patients with signs of MCI. The first study aimed to determine the level of agreement between 'expert' clinicians. The second study aimed to determine the level of agreement between clinicians from Australia and Norway against the 'expert' clinicians and to determine the effect of the level of clinician familiarity with the system on their reliability.

2. Methods

Since this paper reports on two studies, the methods are outlined separately. Fig. 1 provides a flow-chart of the overall study design. The studies were conducted from January 2002 till December 2003. Approval to conduct both studies was obtained from the Curtin University of Technology, Human Research Ethics Committee, Perth, Western Australia.

2.1. Study 1

Patients with NS-CLBP were independently assessed by two 'expert' clinicians and agreement between their diagnoses determined, based on comprehensive subjective and physical examination.

2.1.1. Patients

Patients with a classification of NS-CLBP and MCI seeking physiotherapy treatment were recruited from a private multi-disciplinary orthopaedic clinic in the Perth metropolitan area. After screening and further clinical assessment using strict criteria for inclusion and exclusion (Table 2), 35 patients were selected (17 males and 18 females; mean age 37 ± 12.73 years; duration of LBP 5.6 ± 6.0 years; Revised-Oswestry disability score $37 \pm 11\%$; Body Mass Index $23.1 \pm 2.2\,kg/m^2$). All

Table 1
Different approaches commonly used to classify low back pain (LBP) patients and their limitations

Dimension/category	Approach	Limitation for NS-CLBP population
Uni-dimensional classification systems		
Patho-anatomical	Radiological diagnosis (Bernard and Kirkaldy-Willis, 1987)	Majority (up to 85%) is classified 'non-specific' as no diagnostic imaging procedure is correlated with LBP (Dillingham, 1995; Deyo and Phillips, 1996; Nachemson, 1999; Pearce, 2000) Abnormal findings in asymptomatic individuals are common (Jensen et al., 1994; Boos and Hodler, 1998; Stadnik et al., 1998; Pfirrmann et al., 1999; Borenstein et al., 2001; Humphreys et al., 2002)
	Identify nociceptive source based on diagnostic injections (Bernard and Kirkaldy-Willis, 1987; Bogduk, 1995; Young et al., 2003)	No insight into the underlying mechanism responsible for the LBP disorder (may be driven by neurophysiological, bio-mechanical and/or psycho-social factors)
Signs and Symptoms	'Treatment based' approach, using a cluster of signs and symptoms to classify LBP (Delitto et al., 1995)	For acute LBP only, a similar approach for chronic LBP has not yet been reported Uni-dimensional approach based only on signs and symptoms has limitations as the basis of a mechanism-based CS
Prognosis	Based on the future outcome of the patient (Engel and von Korff MKaton, 1996; Dionne et al., 1997; Krause et al., 1998)	Of limited use for selection of treatment or management Poor prognosis might simply arise because an appropriate treatment addressing the underlying mechanism behind the pain disorder has not yet been identified or tested
Mechanism-based	Hypothesized mechanism behind the disorder is one of disc derangement (McKenzie, 1981)	The validity of this approach for NS-CLBP is limited
	Sahrmann (2001): a classification approach for LBP consisting of five different categories based on signs and symptoms and the premise that 'impairments' in the way people move are the underlying factor of the musculo-skeletal pain and dysfunction	No consideration is given to the stage of the disorder, area of pain, specific versus non-specific LBP, or yellow flags Classification model assumes that all LBP disorders lie within this classification, ignoring the several other dimensions of LBP within the bio-psychosocial framework
Multi-dimensional classification systems		
Stage patho-anatomical signs and symptoms psychosocial	Quebec Task Force Classification (Spitzer, 1987): based on stage of the disorder (acute, sub-acute or chronic), patho-anatomical diagnosis (specific or non-specific, 'red' flags), signs and symptoms (area of pain referral), 'yellow' flags and work status (psycho-social)	Designed to assist in making clinical decisions (e.g. surgery or conservative treatment), establishing a prognosis, and evaluating the quality of care for patients with LBP Of limited use in physiotherapy assessment or treatment planning (Padfield et al., 2002). Does not consider the underlying mechanism behind NS-LBP disorders (apart from differentiating somatic referred from radicular pain)

patients had the protocol explained to them and provided signed consent prior to entering the study.

2.1.2. Examiners

The two examiners were musculo-skeletal physiotherapists. One clinician (PO'S) was the developer of the CS and had 18 years experience with patients with LBP. The other clinician (WD) had 12 years of clinical experience with patients with LBP and extensive training by the developer.

2.1.3. Procedure

Prior to conducting this study, 20 patients participating in a different study conducted by the authors were independently examined and a clinical diagnosis determined by the 'expert' clinicians. The aim of this pilot study was to refine the specific criteria for assignment to each of the five sub-categories and to further train WD (O'Sullivan, 2000, 2004b) (Appendix A). With the aid of videotapes of subjects' postures and movements, the clinicians' diagnoses were compared and discussed, and operational definitions were refined.

Table 2
Inclusion and exclusion criteria for NSLBP patients with Motor Control Impairment[a]

Inclusion criteria
A history of chronic (>3 months) LBP
Pain only located to the lower lumbar spine (L4/5 or L5/S1) with minimal radiation
Absence of impaired movement of the symptomatic segment in the painful direction of movement or loading (based on clinical joint motion palpation examination)
Associated impairments in the control of the motion segment(s) in the provocative movement direction(s)
Clear mechanical basis of disorder: specific postural and functional movements that aggravate and ease symptoms; relief of symptoms when reducing the strain to the symptomatic spinal segment in the provocation direction
Exclusion criteria
More generalized low back pain (beyond L4-5, L5-S1 region) and/or radiating pain
Dominant non-organic features (yellow flags)
Impaired movement of the symptomatic segment in the painful direction of movement or loading (based on clinical joint motion palpation examination)
Based on medical assessment (by general practitioner and/or specialist, including radiological imaging): specific diagnoses for LBP disorder (disc prolapse with radicular pain, severe scoliosis, spondylo-arthrosis, spondylolisthesis, inflammatory or other specific disorders), previous back surgery, and serious causes of LBP (red flag pathology)

[a]All features within the inclusion criteria had to be present; based on O'Sullivan (2000, 2004b).

Following informed consent, patients were allocated to one of two examiners. The order of testing by the two examiners varied but for practical reasons could not be randomized. A full clinical examination was performed by the first examiner to identify patients with NS-CLBP who had a classification of MCI based on strict inclusion and exclusion criteria (Table 1). The comprehensive history of the disorder involved: screening for yellow and red flags, reviewing medical imaging, questioning the patient regarding symptom provocation and relief. The full physical examination consisted of a series of active and functional movement tests, articular tests to determine mobility and level of symptom provocation, neural tissue examination, and tests for spinal motor control (O'Sullivan, 2000, 2004b). Patients were then sub-classified into one of the five patterns as per O'Sullivan (O'Sullivan, 2000, 2004b). Within a maximum of 1 week (most patients were evaluated by the second examiner within 24 h) the second examiner performed a similar full examination and nominated a classification.

The two examiners acted entirely independently and were blind to the other's classification of the disorder. Assessment sheets were placed in sealed opaque envelopes and filed for later analysis. Patients were asked not to provide the second examiner with any information regarding the first examination process.

2.2. Analysis study 1

Kappa-coefficient and %-of-agreement were calculated to determine the level of agreement between the 'expert' clinicians (Portney and Watkins, 2000). The Kappa-coefficient is a reliability statistic that corrects for agreement due to chance (Altman, 1991). Data were analysed using SPSS Version 10.0.

2.3. Study 2

2.3.1. Examiners

Thirteen clinicians (physiotherapists and medical doctors) of two geographically separate regions (seven from Western Australia and six from Norway) were invited to participate based on familiarity with the CS. Examiners' characteristics are displayed in Table 3.

The examiners were classified into two sub-groups based on their level of specific training and clinical experience with the CS. All examiners were required to sign a consent form.

2.3.2. Procedure

Patients who participated in Study 1 were asked to consent to be videotaped and to complete a self-reported pain questionnaire. If consent was obtained, they were videotaped performing a series of postures and functional movements that represented commonly reported aggravating postures and movements of these patients (O'Sullivan, 2000, 2004b). This included usual posture in standing, forward bending and return, backward bending and return, single leg standing, usual sitting posture, slump posture, erect upright posture and sit-to-stand to sit. Thirty patients classified identically by the two 'expert' clinicians (Study 1) gave consent to be videotaped. Of these, 25 patients were randomly selected to fill the previously determined numbers (based on statistical advice) for each of the five patterns. For each of these 25 patients, a case report was written and videotapes were edited in a standard manner.

Approximately 1 month prior to the testing day each participating examiner received an instruction package consisting of a synopsis of the study methodology and a comprehensive summary of the study procedure. Three weeks prior to the testing day a clinical seminar was held

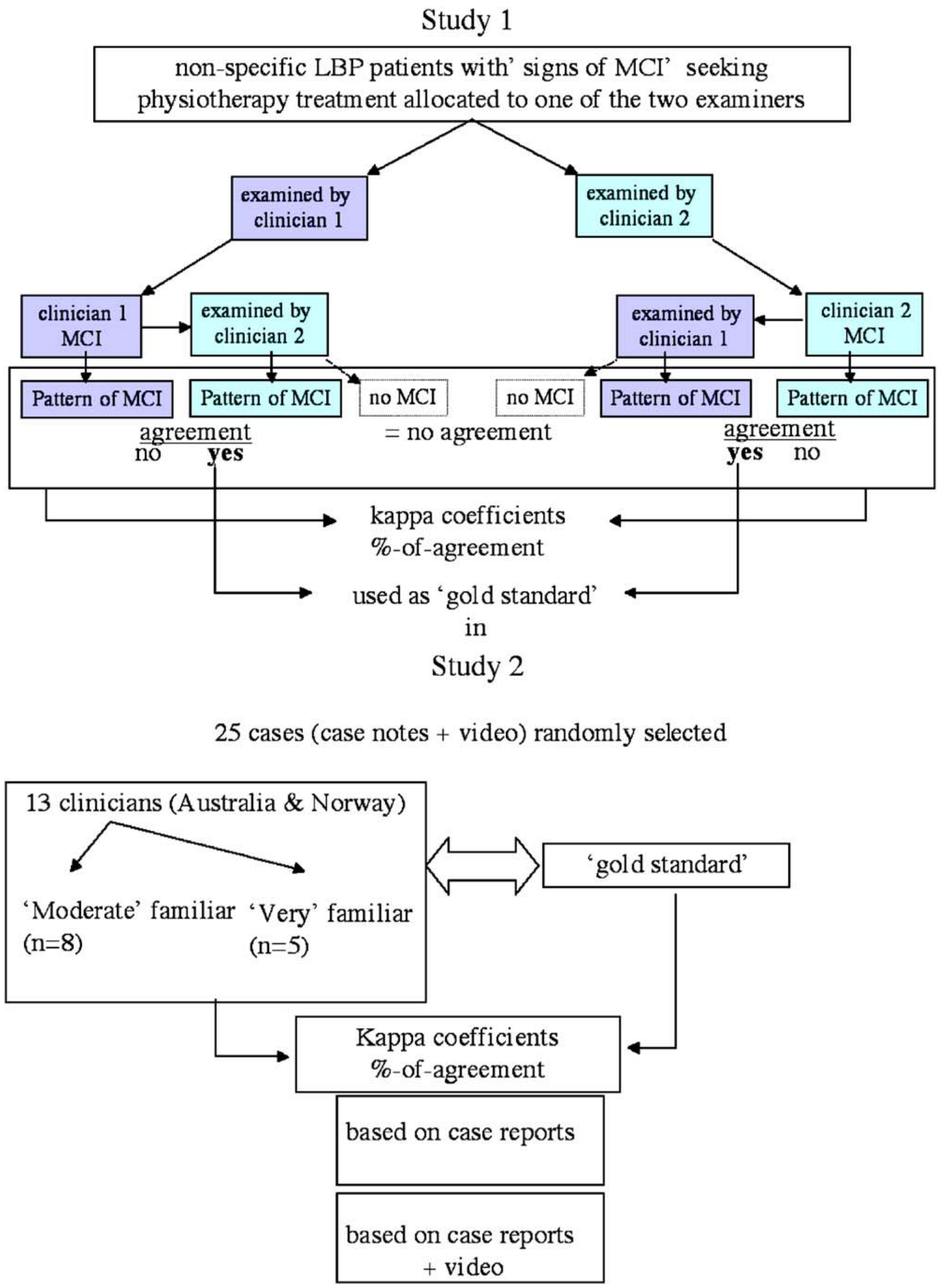

Fig. 1. Flow-chart of study design; MCI = motor control impairment.

by the developer of the CS for all the Western Australian examiners. One week prior to the testing day a revision session was held. For logistical reasons the training for the Norwegian examiners was slightly different. All Norwegian examiners had previously undertaken two clinical workshops based on the CS conducted by the developer. The same instruction package was sent to Norway 3 weeks prior to the testing. A 2-day workshop was held prior to testing.

The blinded examiners had to initially determine the classification for each patient based on the case reports only. In addition, examiners were given the video presentation and were asked to classify the patient based on the combined information. Each examiner placed their assessment booklet in an opaque envelope, which was then sealed prior to further analysis.

2.4. Analysis study 2

Kappa-coefficient and %-of-agreement (Portney and Watkins, 2000) was calculated to determine the level of agreement between the 'gold standard' (as determined

Table 3
Characteristics of examiners (Study 2) as per level of familiarity with Classification System (CS)

	'Moderate' familiar ($n = 8$)	'Very' familiar ($n = 5$)
Specific training in CS	Clinical education sessions and/or workshops regarding the CS with the developer of the CS	Postgraduate training under direct supervision of developer of CS
Age (years)	45 (range: 40–54; SD: 7)	32 (range: 30–33; SD: 1)
Clinical experience (years)	20 (range: 10–29; SD: 7)	9 (range: 7–11; SD: 2)
Speciality	1 GP/physical medicine 1 MD/clinical neurologist 3 Musculo-skeletal physiotherapists 2 Physiotherapists	4 Musculo-skeletal physiotherapists 1 Sports physiotherapist

Table 4
Results of Studies 1 and 2

			Kappa	%-of-agreement
Study 1	'Expert' clinicians ($n = 2$)	SE + PE	0.96	97%
Study 2	All clinicians ($n = 13$)	S	0.32 (0.13–0.54)	48% (32%–64%).
		S + V	0.61 (0.47–0.80)	70% (60%–84%)
	'Moderate familiar' clinicians ($n = 8$)	S	0.28 (0.13–0.37)	44% (32%–52%)
		S + V	0.55 (0.47–0.64)	65% (60%–72%)
	'Very familiar' clinicians ($n = 5$)	S	0.40 (0.39–0.54)	54% (48%–64%)
		S + V	0.71 (0.58–0.80)	78% (68%–84%)

Average Kappa-scores (range) and average %-of-agreement (range) SE + PE = based on a comprehensive subjective and physical examination, S = based on subjective information only (case notes), S + V = based on subjective and video. Guidelines for interpreting the strength of the Kappa statistic: <0: poor; 0.00–0.20: slight; 0.21–0.40: fair; 0.41–0.60: moderate; 0.61–0.80: substantial; 0.81–1.00: almost perfect (based on Altman, 1991).

by the 'expert' clinicians) and the other examiners. Agreement was also analysed based on the level of familiarity with the CS, based on subjective information only, and based on subjective information plus the videotaped recordings. Descriptive statistics were used for the analysis of correct classification for each pattern. Data were analysed using SPSS Version 10.0.

3. Results

3.1. Study 1

Based on independent patient examinations 'expert' clinicians demonstrated almost perfect agreement (Kappa-coefficient 0.96; %-of-agreement 97%) (Table 4).

3.2. Study 2

The agreement between examiners and 'expert' clinicians based on subjective information and video was substantial (Table 3). Agreement was reduced when examiners made a classification decision based only on subjective information, and among those examiners who had less familiarization with the CS (Table 3). Fig. 2a–e shows the correct classification (%) by all examiners for each pattern. All five patterns could be reliably identified, with the Flexion Shifting pattern best identified (82%), and the Active Extension pattern least correctly identified (62%).

4. Discussion

The objective of Studies 1 and 2 was to assess the inter-examiner reliability of a CS for NS-CLBP with MCI as proposed by O'Sullivan (2000, 2004b). Results of Study 1 revealed that there was almost perfect agreement between 'expert' clinicians, in identifying and classifying patients with NS-CLBP into specific subgroups of MCI based on a comprehensive clinical examination (Table 4). Results of Study 2 indicate substantial clinical agreement across all five patterns based on combined subjective case reports and video observation of postures and movements. Good inter-examiner reliability is an essential first step for a CS to be valid and to be of use, clinically and in research settings (Delitto et al., 1995).

The poor reliability when classification was based only on subjective findings was expected. This finding supports that the CS is highly dependent on the assimilation of both the subjective and physical examination (O'Sullivan, 2000, 2004b).

Study 2 also aimed to evaluate the importance of specific training in the CS, relative to the ability to accurately apply the CS. The results of this study

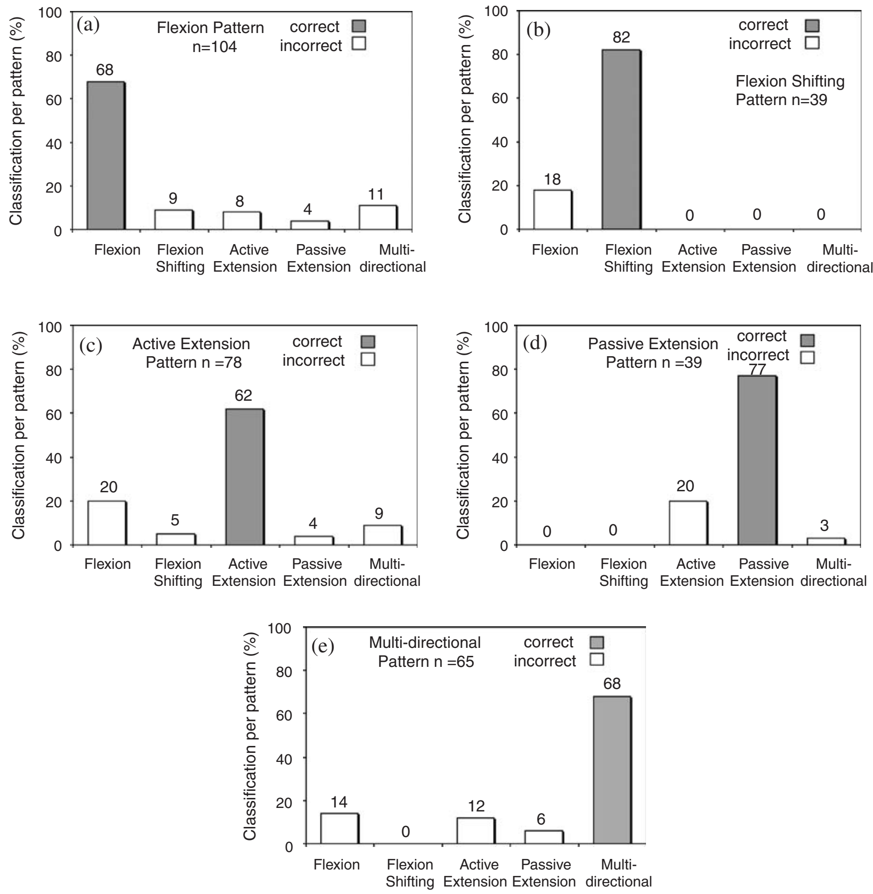

Fig. 2. (a–e) Classification per different pattern (in %) by all examiners in Study 2; n = total number of that specific pattern included × 13 (total number of examiners).

(Table 4) show a very clear pattern of improved reliability associated with more specific (postgraduate) training. This finding is consistent with Strender et al. (1997) who state that the amount of formal instruction (i.e. continuing education) and specific clinical experience in examination procedures and classification rules is a necessary prerequisite to improving reliability.

There appears to be a special need for a mechanism-based CS for NS-CLBP based on a bio-pyscho-social framework (Woolf et al., 1998; McCarthy et al., 2004; O'Sullivan, 2004a; O'Sullivan, 2004b). It is acknowledged that to validate this novel CS as a mechanism-based CS, a multi-step process is required and cannot be based solely on inter-examiner reliability. For this reason a model for clinical research into classification of NS-CLBP has been proposed by the authors (Dankaerts et al., 2004; O'Sullivan, 2004a). This model consists of different stages, each stage dealing with different criteria. Fig. 3 presents a flow-chart summary of the model.

This multi-dimensional mechanism-based CS is not an alternative to existing CSs but can be seen as a new development for a sub-group within the NS-CLBP, integrating different aspects of established CSs. For example, the CS proposed by O'Sullivan fits within the QTFC as it uses several criteria set forward by the QTFC: the patient sample consists of 'non-specific', 'chronic', 'LBP patients without radiation below the gluteal folds', absence of 'red and dominant yellow flags' and absence of 'neurological signs'. The proposed CS by O'Sullivan can be seen as a further sub-classification of Category 1 of the QTFC. Similar to McKenzie's method (McKenzie, 1981; Donelson, 2001), O'Sullivan's CS is based on a comprehensive patient assessment. Like Delitto's (1995) treatment-based CS or McKenzie's method (McKenzie, 1981), O'Sullivan links a very specific intervention to each of the five patterns.

Of particular interest for direct comparison with the proposed multi-dimensional CS is the CS developed by Sahrmann (2001) which appears to be more uni-dimensional in nature. Both classification models assume MCIs as a possible underlying factor in LBP disorders. But there are some substantial differences in the proposed ways to validate the CS and in the method used to classify the patients. Van Dillen et al. (1998) investigated the reliability (among trained therapists) of the individual tests used in criteria for classification according to Sahrmann (2001). The nature of their study

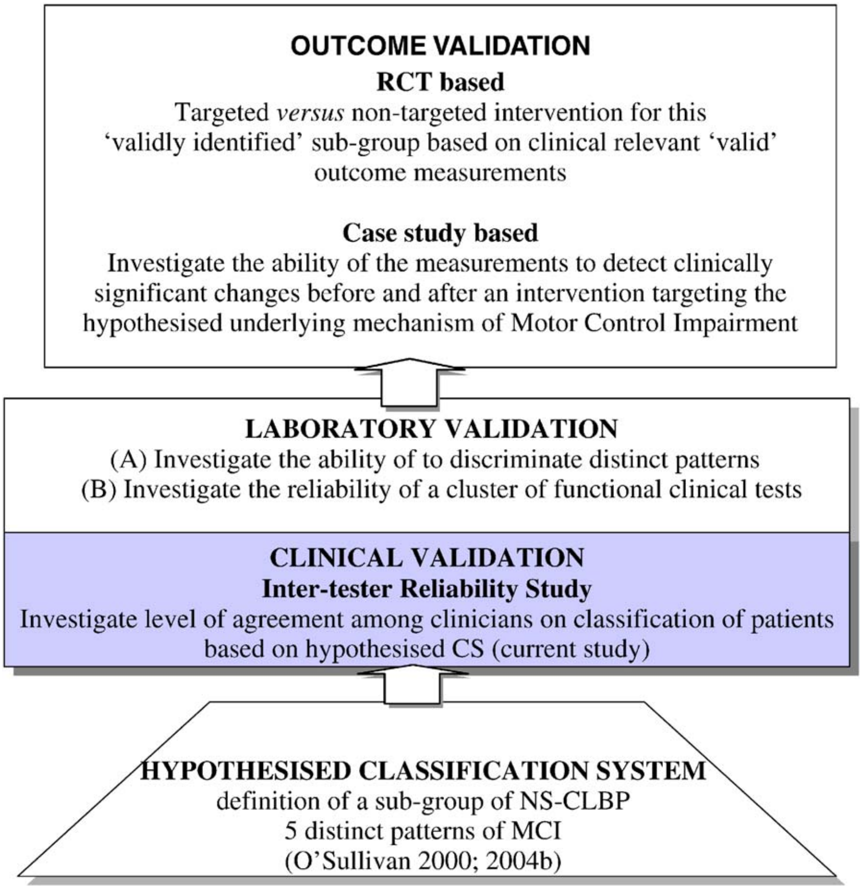

Fig. 3. Flow-chart of proposed new model for validation of mechanism-based classification system (CS) for NS-CLBP with Motor Control Impairment (MCI). Current study shaded.

design assessed reliability based on individual physical examination items and did not give any insight into the ability of the clinicians to classify the patients into the proposed categories. This reliability study served as a pilot study for a validation study on the CS proposed by Sahrmann (Van Dillen et al., 2003b). According to Bailey (1994), a principal aim of classification is ordering entities into groups with maximum between group heterogeneity and within group homogeneity. In Van Dillen et al. (2003b), the location of symptoms varied from low back only to all kinds of referred locations. All three different stages of LBP disorder were included, without consideration for patho-anatomical findings nor the presence of non-organic signs. The fact that these inclusion criteria consist of several different QTFC categories might have led to a heterogeneous sample.

Rather than relying purely on signs and symptoms (Van Dillen et al., 1998, 2003a, b), the current studies used a process of diagnostics to make a clinical determination as to whether the MCI is the driving mechanism behind the disorder or is simply a secondary effect of another process (O'Sullivan, 2004a, b). This process of diagnostics is described in detail elsewhere (Elvey and O'Sullivan, 2004). In contrast with Van Dillen et al. (1998, 2003b), the authors of the current study place a strong emphasis on the subjective history and pain behaviour. Within the CS it is critical in interpreting how the symptoms (as described by the patient during subjective examination) are influenced by changes in postural alignment and movement. Another purpose of taking and integrating history findings is to determine the presence of dominant non-organic features. Strong evidence exists to suggest that psychosocial factors can be an important component of certain NS-LBP disorders (Linton, 2000). Psychological processes (cognition, stress, fear, anxiety and depression) are also known to alter motor behaviour (Hodges and Moseley, 2003) influencing patient's posture and movement (Hodges and Moseley, 2003). Attempts to 'normalize' the movement or MCI in many of these disorders would be inappropriate and ineffective (Elvey and O'Sullivan, 2004; O'Sullivan, 2004a). In Van Dillen et al. (1998, 2003b), no other physical examination was performed to identify other underlying mechanisms of pain response. In contrast, in the current study the patients were firstly identified as having an MCI based on a set of characteristics (Table 2). We found it essential to include 'joint motion palpation' in Study 1. Firstly, because it was deemed important to identify whether the observed control impairment was linked to the symptomatic level of the patient. Secondly, to identify if the pain disorder is linked to an impairment of 'movement' or 'control' (Elvey and O'Sullivan, 2004). According to the present authors, in the case of a painful impairment of movement, a treatment (such as manipulative techniques) aiming to promote movement into the painful range is the treatment of choice (Elvey and O'Sullivan, 2004).

4.1. Limitations and recommendations for further studies

A limitation of this study was the fact that only the clinicians in Study 1 had to agree on identifying MCI patients from the larger LBP population. Patients in Study 2 must therefore be seen as a selected group that does not fully represent the general population of NS-CLBP patients which may be more difficult to classify. Further studies are required to test the ability of clinicians to identify patients with MCI within this NS-CLBP population.

The use of 'expert' clinicians' classification as 'gold standard' is another limitation of this study. But in the absence of a true criterion standard for MCI diagnosis, this method was justified and has been used by others (Gracovetsky et al., 1995).

Because of practical and logistic issues, it was decided to use videotaped recordings of the patients in Study 2. Videotaping has been previously used during reliability studies on visual analysis of gait (Krebs et al., 1985; Eastlack et al., 1991), scapular dysfunction (Kibler et al., 2002) and spinal movements (Fritz et al., 2000). Videotaping has been recommended as an alternative to a test–retest design for assessing inter-examiner reliability with patients who have LBP (Delitto et al., 1992). Both limitations and advantages of videotaping are recognized by the authors. The lack of an actual clinical examination for Study 2 is a limitation and may account for some of the discrepancy in the results between Studies 1 and 2. The advantages of videotaping include: not having to place undue stress on the patient, potentially altering his or her clinical status, while allowing a greater number of participants from geographically distinct regions.

In spite of these limitations, the current results support good to high level-of-agreement across all categories based on the methods used.

As mentioned above, a multi-step process is required to validate this novel multi-dimensional CS as a mechanism-based classification model. Based on the new proposed model (Fig. 3), several studies have been undertaken to add laboratory and outcome validity to the CS. A laboratory-based test battery, including EMG and 3D-motion analysis, is currently being employed to further validate the clinical diagnosis, determine motor control differences in pain sub-groups with normative data, and ultimately, provide outcome measures for specific interventions.

5. Conclusions

The main aim of these two studies was to investigate the inter-examiner reliability of a CS for NS-CLBP patients with MCI. Substantial to excellent reliability was found depending on the level of familiarity. Further research is required to further validate the proposed CS as a mechanism-based CS. The authors believe that the acceptance and integration of a multi-dimensional mechanism-based CS for NS-CLBP with MCI could have profound implications leading to the application of specific 'targeted' interventions for identified sub-groups, and subsequently enhanced treatment efficacy as suggested by Leboeuf-Yde et al. (1997, 2001).

Acknowledgment

This study was carried out whilst the first author (WD) was an International Postgraduate Research Scholar in Australia and was supported financially by the Head of School of Physiotherapy Scholarship, Curtin University of Technology Western Australia.

The authors would also like to acknowledge all examiners (CP, DB, LP, LG, JW, RY, MK, JS, KH, LL, KO, KF, AT) who participated in this study. Special thanks to the administrative staff at 'Body Logic' for their help in co-ordinating the appointments for patients, Dr. Ritu Gupta for her statistical advice and Marina Wise for the linguistic corrections.

Appendix A

The different sub-groups of MCI Patterns and their clinical presentation are briefly described below. Based on O'Sullivan (2000, 2004b)

Flexion pattern

Definition: MCI of the lumbar spine with a tendency to flexion strain (loss of segmental lordosis) at the symptomatic segment. Flexion pain disorders are associated with functional loss of motor control into flexion resulting in an excessive abnormal flexion strain.
Provocative postures/activities: all flexion-related postures (e.g. slouched sitting) and functional activities

(forward bending, cycling) are commonly reported as being painful.
Easing postures/activities: extension postures/activities where the lumbar spine is lordosed (e.g. standing, sitting with a lumbar roll, walking).
Posture and movement analysis: tendency to present with a loss of lumbar lordosis during sitting and standing postures. The pelvis is often positioned in posterior pelvic tilt. During all functional tasks the same tendency to have a loss of lordosis at the 'symptomatic level' is noted. Forward bending movements commonly reveal a tendency of an early 'loss of lower lumbar lordosis' (lumbar curve reversal). Similar loss of lordosis is accentuated in other functional tasks like sit-to-stand, squatting and gait. This is associated with an increased lordosis in the upper lumbar and lower thoracic spine.
Specific posture and movement control tests: inability/lack of motor control to anterior rotate pelvis and extend lower lumbar spine independent from thorax during above-mentioned aggravating postures/movements.

Flexion/lateral shifting pattern

Definition: MCI around the lumbar spine with a tendency to flex and laterally shift at the symptomatic segment.
Provocative postures/activities: reaching and rotating in one direction in association with flexion postures and / or movements.
Easing postures/activities: relief in extended or lordotic postures, stretching to the opposite side from the shift, shift correction (contra-lateral glide from pelvis).
Posture and movement analysis: similar to the flexion pattern there is a loss of lumbar segmental lordosis at the affected level with the key feature here an associated lateral shift at the lower lumbar spine level. Minimal precipitation of their spine might deviate into a lateral shift position. E.g.: the lateral shift is accentuated when standing on the foot ipsi-lateral to the shift. Sagittal spinal movements reveal a tendency to laterally deviate during flexion and this is commonly associated with an arc of pain. Tests like 'sit to stand' usually reveal a typical flexion pattern presentation (see above) plus a tendency towards lateral trunk shift during the movement with increased weight bearing on the lower limb on the side of the shift.
Specific posture and movement control tests: inability/lack of motor control to anterior rotate pelvis and extend lower lumbar spine independent from thorax during above-mentioned aggravating postures/movements with an associated lateral deviation

Active extension pattern

Definition: MCI around the lumbar spine with a tendency to hold the lumbar spine actively into extension.
Provocative postures/activities: all extension-related postures (standing, erect sitting) and functional activities (carrying out overhead activities, fast walking, running and swimming) are commonly reported as being painful. Also commonly reported as a provocative activity is forward bending (with the key feature here being the tendency to hold the lumbar spine into segmental hyperextension).
Easing postures/activities: flexion postures/activities where the lumbar spine is flexed (e.g. crook lying, slouched sitting).
Posture and movement analysis: tendency for the lumbar spine to be actively held into segmental hyper-lordosis at the symptomatic segment during upright sitting and standing postures. During all functional tasks such as sit to stand, squatting and forward bending the same tendency to hyper-lordose at the 'symptomatic segment' is noted. Forward bending movements commonly reveal increased hip flexion and a tendency of a late 'loss of lordosis' (beyond mid range of flexion) or no lumbar curve reversal. Return to neutral from a forward bended position reveals an early hyper-lordosing of the spine at the symptomatic segment.
Specific posture and movement control tests: inability/lack of motor control to initiate a posterior pelvic during above-mentioned aggravating postures/movements.

Passive extension pattern

Definition: MCI around the lumbar spine with a tendency to passively over-extend at the symptomatic segment of the lumbar spine.
Provocative postures/activities: similar to the active extension pattern all extension-related postures (standing, erect sitting) and functional activities (carrying out overhead activities, fast walking, running and swimming) are commonly reported as being painful.
Easing postures/activities: flexion postures/activities where the lumbar spine is de-lordosed (e.g. crook lying, slouched sitting).
Posture and movement analysis: tendency for patients to stand into a sway-back posture (thorax posterior to the pelvis) with a segmental hinging at the symptomatic level. Forward bending is often pain free, but on return to neutral they tend to over-extend at the symptomatic level (hinge into extension) and sway pelvis anterior.

Specific posture and movement control tests: inability/lack of motor control to extend the thoraco-lumbar spine above the symptomatic segment with a tendency to hinge into extension at this segment.

Multi-directional pattern

Definition: multi-directional MCI around the lumbar spine

Provocative postures/activities: multi-directional nature of this pattern often reveals pain all weight bearing postures and functional activities.

Easing postures/activities: difficulty to find relieving positions during weight bearing

Posture and movement analysis: patient may assume a flexed, extended or laterally shifted spinal posture, and may frequently have to alternate them. Excessive segmental shifting and hinging may be observed in all directions, with associated 'jerky' movement patterns and reports of 'stabbing' pain on movement in all directions with observable lumbar erector spinae muscle spasm.

Specific posture and movement control tests: patients have great difficulty assuming neutral lordotic spinal postures, with over shooting into flexion, extension or lateral shifting postures.

References

Altman DG. Practical statistics for medical research. London: Chapman & Hall; 1991.

Bailey KD. Typologies and taxonomies: an introduction to classification techniques. Thousand Oaks, California: Sage Publications; 1994.

Bernard TN, Kirkaldy-Willis WH. Recognizing specific characteristics of nonspecific low back pain. Clinical Orthopedics 1987;217:266–80.

Biedermann HJ, Shanks GL, Forrest WJ, Inglis J. Power spectrum analyses of electromyographic activity. Discriminators in the differential assessment of patients with chronic low-back pain. Spine 1991;16(10):1179–84.

Binkley J, Finch E, Hall J, Black T, Gowland C. Diagnostic classification of patients with low back pain: report on a survey of physical therapy experts. Physical Therapy 1993;73(3):138–50.

Bogduk N. The anatomical basis for spinal pain syndromes. Journal of Manipulative and Physiological Therapeutics 1995;18(9):603–5.

Boos N, Hodler J. What help and what confusion can imaging provide? Baillieres Clinical Rheumatology 1998;12(1):115–39.

Borenstein D.G., O'Mara J.W., Boden S.D., Lauerman W.C., Jacobson A., Platenberg C., Schellinger D., Wiesel S.W., The value of magnetic resonance imaging of the lumbar spine to predict low-back pain in asymptomatic subjects: a seven-year follow-up study. Journal of Bone and Joint Surgery 2001;83(9)A:1306–11.

Borkan JM, Koes B, Reis S, Cherkin DC. A report from the Second International Forum for Primary Care Research on Low Back Pain. Reexamining priorities. Spine 1998;23(18):1992–6.

Borkan J, Van Tulder M, Reis S, Schoene ML, Croft P, Hermoni D. Advances in the field of low back pain in primary care: a report from the fourth international forum. Spine 2002;27(5):E128–32.

Bouter LM, van Tulder MW, Koes BW. Methodologic issues in low back pain research in primary care. Spine 1998;23(18):2014–20.

Burton AK, McClune TD, Clarke RD, Main CJ. Long-term follow-up of patients with low back pain attending for manipulative care: outcomes and predictors. Manual Therapy 2004;9(1):30–5.

Croft PR, Macfarlane GJ, Papageorgiou AC, Thomas E, Silman AJ. Outcome of low back pain in general practice: a prospective study. British Medical Journal 1998;316(7141):1356–9.

Dankaerts W, O'Sullivan P, Burnett A, Straker L. What precedes RCT's? A new model for clinical research into NSCLBP. Eighth International Federation of Orthopaedic Manipulative Therapists' Conference, Cape Town, South Africa, 2004.

Delitto A, Shulman AD, Rose MJ, Strube MJ, Erhard RE, Bowling RW. Reliability of a clinical examination to classify patients with low back syndrome. Physical Therapy Practice 1992;1:1–9.

Delitto A, Erhard RE, Bowling RW. A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative treatment. Physical Therapy 1995;75(6):470–89.

Deyo RA, Phillips WR. Low back pain. A primary care challenge. Spine 1996;21(24):2826–32.

Dillingham T. Evaluation and management of low back pain: an overview. State of the Art Reviews 1995;9(3):559–74.

Dionne CE, Koepsell TD, Von Korff M, Deyo RA, Barlow WE, Checkoway H. Predicting long-term functional limitations among back pain patients in primary care settings. Journal of Clinical Epidemiology 1997;50(1):31–43.

Donelson R, editor. Letter to the Editor Spine 2001;26(16):1827–9.

Eastlack ME, Arvidson J, Snyder-Mackler L, Danoff JV, McGarvey CL. Interrater reliability of videotaped observational gait-analysis assessments. Physical Therapy 1991;71(6):465–72.

Ehrlich GE. Back pain. Journal of Rheumatology 2003(Suppl 67):26–31.

Elvey R, O'Sullivan P. A contemporary approach to manual therapy. Grieve's Modern Manual Therapy. Boyling and Jull, Elsevier 2004:471–94.

Engel CC, von Korff MKaton WJ. Back pain in primary care: predictors of high health-care costs. Pain 1996;65(2-3):197–204.

Ford J, Story I, McKeenen J. A systematic review on methodology of classification system research for low back pain. Musculoskeletal Physiotherapy Australia 13th Biennial Conference, Sydney, Australia, 2003.

Fritz JM, Delitto A, Vignovic M, Busse RG. Interrater reliability of judgments of the centralization phenomenon and status change during movement testing in patients with low back pain. Archives of Physical Medicine and Rehabilitation 2000;81(1):57–61.

Fritz JM, Delitto A, Erhard RE. Comparison of classification-based physical therapy with therapy based on clinical practice guidelines for patients with acute low back pain: a randomized clinical trial. Spine 2003;28(13):1363–72.

Gracovetsky S, Newman N, Pawlowsky M, Lanzo V, Davey B, Robinson L. A database for estimating normal spinal motion derived from non invasive measurements. Spine 1995;20(9):1036–46.

Hodges P, Richardson C. Contraction of the abdominal muscles associated with movement of the lower limb. Physical Therapy 1997;77(2):132–43.

Hodges PW. Changes in motor planning of feedforward postural responses of the trunk muscles in low back pain. Experimental Brain Research 2001;141(2):261–6.

Hodges PW, Moseley GL. Pain and motor control of the lumbopelvic region: effect and possible mechanisms. Journal of Electromyography and Kinesiology 2003;13(4):361–70.

Hodges PW, Richardson CA. Delayed postural contraction of transversus abdominis in low back pain associated with movement of the lower limb. Journal of Spinal Disorders 1998;11(1):46–56.

Hodges PW, Richardson CA. Altered trunk muscle recruitment in people with low back pain with upper limb movement at different speeds. Archives of Physical Medicine and Rehabilitation 1999;80(9):1005–12.

Humphreys SC, Eck JC, Hodges SD. Neuroimaging in low back pain. American Family Physician 2002;65(11):2299–306.

Jensen MC, Kelly AP, Brant-Zawadzki MN. MRI of degenerative disease of the lumbar spine. Magnetic Resonance Quarterly 1994;10(3):173–90.

Katz WA. Musculoskeletal pain and its socioeconomic implications. Clinical Rheumatology 2002;21(Suppl 1):S2–4.

Kibler WB, Uhl TL, Maddux JW, Brooks PV, Zeller B, McMullen J. Qualitative clinical evaluation of scapular dysfunction: a reliability study. Journal of Shoulder and Elbow Surgery 2002;11(6): 550–6.

Krause N, Ragland DR, Fisher JM, Syme SL. Psychosocial job factors, physical workload, and incidence of work-related spinal injury: a 5-year prospective study of urban transit operators. Spine 1998;23(23):2507–16.

Krebs DE, Edelstein JE, Fishman S. Reliability of observational kinematic gait analysis. Physical Therapy 1985;65(7):1027–33.

Leboeuf-Yde C, Manniche C. Low back pain: time to get off the treadmill. Journal of Manipulative and Physiological Therapeutics 2001;24(1):63–6.

Leboeuf-Yde C, Lauritsen JM, Lauritzen T. Why has the search for causes of low back pain largely been nonconclusive? Spine 1997;22(8):877–81.

Leclere H, Beaulieu MD, Bordage G, Sindon A, Couillard M. Why are clinical problems difficult? General practitioners' opinions concerning 24 clinical problems. Canadian Medical Association Journal 1990;143(12):1305–15.

Linton SJA. review of psychological risk factors in back and neck pain. Spine 2000;25(9):1148–56.

Luoto S, Taimela S, Hurri H, Alaranta H. Mechanisms explaining the association between low back trouble and deficits in information processing. A controlled study with follow-up. Spine 1999;24(3):255–61.

McCarthy CJ, Arnall FF, Strimpakos N, Freemont A, Oldham JA. The biopsychosocial classification of non-specific low back pain: a systematic review. Physical Therapy Reviews 2004;9:17–30.

McKenzie R. The lumbar spine. mechanical diagnosis and therapy. Waikanae, New Zealand: Spinal Publications; 1981.

Nachemson A. Back pain: delimiting the problem in the next millennium. International Journal of Law and Psychiatry 1999;22(5–6):473–90.

O'Sullivan PB. Lumbar segmental 'instability': clinical presentation and specific stabilizing exercise management. Manual Therapy 2000;5(1):2–12.

O'Sullivan P. Invited commentary. Journal of Orthopaedic and Sports Physical Therapy 2004a;34(3):109–10.

O'Sullivan P. Clinical instability of the lumbar spine: its pathological basis, diagnosis and conservative management. In: Boyling JD, Jull GA, editors. Modern Manual Therapy. Elsevier 2004b;1: 311–22.

O'Sullivan PB, Twomey LT, Allison GT. Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylolysis or spondylolisthesis. Spine 1997;22(24):2959–67.

O'Sullivan PB, Twomey L, Allison GT. Altered abdominal muscle recruitment in patients with chronic back pain following a specific exercise intervention. Journal of Orthopaedic and Sports Physical Therapy 1998;27(2):114–24.

Padfield B, Chesworth B, Butler R. Use of an outcome measurement system to answer a clinical question: is the Quebec task force classification system useful in an outpatient setting? Physiotherapy Canada Fall 2002:254–60.

Pearce JM. Aspects of the failed back syndrome: role of litigation. Spinal Cord 2000;38(2):63–70.

Pfirrmann CW, Hodler J, Boos N. Diagnostic assessment in lumbar back pain. II. Imaging and image-guided infiltrations. Schweiz Rundsch Med Prax 1999;88(8):315–21.

Portney LG, Watkins MP. Foundations of clinical research: applications to practice. Upper Saddle River, NJ: Pentice Hall; 2000.

Riddle DL. Classification and low back pain: a review of the literature and critical analysis of selected systems. Physical Therapy 1998;78(7):708–37.

Sahrmann SA. Diagnosis and treatment of movement impairment syndromes. Mosby: St. Louis; 2001.

Spitzer WO. Scientific approach to the assessment and management of activity related spinal disorders. Spine 1987;7S:S1–S55.

Spitzer WO, Skovron ML, Salmi LR, Cassidy JD, Duranceau J, Suissa S, Zeiss E. Scientific monograph of the Quebec Task Force on Whiplash-Associated Disorders: redefining "whiplash" and its management. Spine 1995;20(8 Suppl):1S–73S.

Stadnik TW, Lee RR, Coen HL, Neirynck EC, Buisseret TS, Osteaux MJ. Annular tears and disk herniation: prevalence and contrast enhancement on MR images in the absence of low back pain or sciatica. Radiology 1998;206(1):49–55.

Strender LE, Sjoblom A, Sundell K, Ludwig R, Taube A. Interexaminer reliability in physical examination of patients with low back pain. Spine 1997;22(7):814–20.

Van Dillen LR, Sahrmann SA, Norton BJ, Caldwell CA, Fleming DA, McDonnell MK, Woolsey NB. Reliability of physical examination items used for classification of patients with low back pain. Physical Therapy 1998;78(9):979–88.

Van Dillen LR, Sahrmann SA, Norton BJ, Caldwell CA, McDonnell MK, Bloom N. The effect of modifying patient-preferred spinal movement and alignment during symptom testing in patients with low back pain: a preliminary report. Archives of Physical Medicine and Rehabilitation 2003a;84(3):313–22.

Van Dillen LR, Sahrmann SA, Norton BJ, Caldwell CA, McDonnell MK, Bloom NJ. Movement system impairment-based categories for low back pain: stage 1 validation. Journal of Orthopaedic & Sports Physical Therapy 2003b;33(3):126–42.

van Tulder M, Koes B, Bombardier C. Low back pain. Best Practice & Research in Clinical Rheumatology 2002;16(5):761–75.

Waddell G. 1987 Volvo award in clinical sciences. A new clinical model for the treatment of low-back pain. Spine 1987;12(7):632–44.

Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. Bulletin of the World Health Organanization 2003;81(9):646–56.

Woolf CJ, Bennett GJ, Doherty M, Dubner R, Kidd B, Koltzenburg M, Lipton R, Loeser JD, Payne R, Torebjork E. Towards a mechanism-based classification of pain? Pain 1998;77(3):227–9.

Young S, Aprill C, Laslett M. Correlation of clinical examination characteristics with three sources of chronic low back pain. Spine Journal 2003;3(6):460–5.

Zimny NJ. Diagnostic classification and orthopaedic physical therapy practice: what we can learn from medicine. Journal of Orthopaedic & Sports Physical Therapy 2004;34(3):105–9 discussion 110-5.