Interrater Reliability of a New Classification System for Patients with Neural Low Back-Related Leg Pain

Axel Schäfer, PT[1,2]; **Toby M. Hall**, MSc[3,4]; **Kerstin Lüdtke**, MSc[2];
Joachim Mallwitz, MD[5]; **Noelle K. Briffa**, PhD[6]

In 1997 the importance of identifying subgroups of low back pain (LBP) was given top ranking by the Second International Forum for Primary Care Research on Low Back Pain[1]. Classification of "non-specific" LBP patients into distinct subgroups is essential to deliver specific and effective treatment in primary care and to determine which treatments will be most effective in clinical research. Ten years later, the same issue is still seen as the number-one research priority by primary care practitioners[2]. There is considerable controversy in the literature regarding correct diagnosis and classification[3,4], in particular for low back-related leg pain, which is associated with LBP in up to 65% of cases[5-8].

ABSTRACT: The aim of this study was to investigate the reliability of a new classification system for low back-related leg pain arising from neural tissue dysfunction. Leg pain is a frequent accompaniment to back pain and is an indicator of the severity and prognosis of the disorder. For optimal patient care, treatment should be directed according to the identified pathophysiological mechanisms. The authors have proposed a sub-classification of neural low back-related leg pain into four categories, each requiring a different management strategy: Central Sensitization (CS), comprising major features of sensitization of the somatosensory system; Denervation (D), arising from significant axonal compromise without evidence of major central nervous system changes; Peripheral Nerve Sensitization (PNS), arising from nerve trunk inflammation without clinical evidence of significant denervation; and Musculoskeletal pain (M), referred from non-neural structures such as the disc or facet joints. The purpose of this study was to investigate the interrater reliability of this classification system. Forty consecutive patients with unilateral low back-related leg pain were independently assessed by five pairs of examiners using a physical examination protocol, screening for central sensitization of the somatosensory system, neurological deficit, and nerve tissue mechano-sensitization. Subjects were classified as follows: CS 30%, D 27.5%, PNS 10%, and M 32.5%. Interrater reliability was good with 80% agreement and a κ of 0.72 (95% Confidence Interval (CI) .57–.86). The findings of the study demonstrate that patients with low back-related leg pain can be reliably classified to one of the four proposed groups.

KEYWORDS: Classification, Interrater Reliability, Leg Pain, Low Back Pain

Traditionally, a distinction has been made between radicular pain and somatic referred pain[9], but in the absence of a true diagnostic gold standard, the differentiation between the two remains difficult[10]. Failure to correctly classify subjects into homogenous subgroups may be one explanation for inconsistent findings of randomized controlled trials investigating the effectiveness of conservative treatment of patients with sciatica[11,12]. This is in spite of the fact that a number of classification systems for LBP incorporating radiating leg pain have been published since 1981[13-15]. Current studies have demonstrated that conservative therapy, when implementing classification systems, can be effective for two subgroups of patients with LBP and lower limb symptoms[16-19]. Patients with low back-related leg pain without neurological signs, whose symptoms centralize with repeated movement in a specific direction (directional preference), have shown to improve with direction-specific treatment[16,18,19]. Furthermore, patients with lower limb symptoms, signs of nerve root compression, and either a crossed

[1]Professorship for Physiotherapy, HAWK—University of Applied Sciences and Arts, Hildesheim, Germany
[2]PT, Rückenzentrum am Michel, Hamburg, Germany
[3]Director, Manual Concepts, Perth, Australia
[4]Senior Lecturer, Curtin University of Technology, Perth, Australia
[5]MD, CEO Rückenzentrum am Michel, Hamburg, Germany
[6]Assist. Professor, Curtin University of Technology, Perth, Australia

straight leg raise or peripheralization of symptoms with extension movements appear to benefit from a traction-based treatment intervention[17].

The categorization of these two groups of patients with low back-related leg pain arises from a treatment-based classification system for LBP[14]. However, there is ample evidence in the literature suggesting the existence of at least two further subgroups with similar symptoms but different patho-mechanisms. Hence, our classification system can be seen as an extension of existing classification systems by adding one further dimension for neural-related pain. One subgroup consists of patients with pain arising from a nerve lesion affecting the somatosensory system, resulting in sensory hypersensitivity, closely connected to central pain mechanisms[20,21]. One other subgroup consists of patients with mechano-sensitization of the nerve trunk due to inflammatory processes in and around the peripheral nerve[22-24].

It has been suggested that for the more complex pain conditions related to nerve injury, a classification based on patho-mechanisms may be more suitable[25,26] as it offers a rational approach to treatment and greater diagnostic sensitivity, and it has the potential to provide information about the prognosis and natural course of the disorder[26]. Therefore, we developed a patho-mechanism-based classification system for low back-related leg pain[27], extending the ideas originally presented by Elvey and Hall[28]. The purpose of refining this classification system was to improve treatment outcome, particularly with respect to identifying patients most likely to respond to neural mobilization. Depending on the assumed predominance of patho-mechanisms, low back-related leg pain is differentiated into four distinct categories. Classified hierarchically, these categories are Central Sensitization (CS), comprising major features of sensory sensitization; Denervation (D), arising from significant axonal compromise without major evidence of the above; Peripheral Nerve Sensitization (PNS), arising from nerve trunk inflammation without clinical evidence of significant denervation; and Musculoskeletal pain (M), referred from non-neural structures such as the lumbar intervertebral disc or zygapophyseal joints[27] (Figure 1).

The category PNS has been previously described[28,29] and is thought to characterize patients who may respond to manual therapy involving neural mobilization. There is reasonable evidence showing that patients with signs and symptoms corresponding to PNS benefit from a neural mobilization treatment in the upper quarter[30-33] and limited evidence for a response in the lower quarter[34,35]. In contrast, patients classified as D might need a different treatment approach as it seems unlikely that structural nerve damage will respond favorably to neural mobilization. The effects of neural mobilization treatment in a patient with lumbar radiculopathy with signs and symptoms corresponding to group D have been investigated in a case report; no treatment benefits were demonstrated[36]. Additionally, the effects of a neural mobilization intervention on pain and disability have been investigated in patients following decompression surgery[37]. Patients referred for spinal decompression surgery are likely to have significant evidence of nerve root compression, with corresponding sensory and motor deficit, consistent with group D. The results of that study indicated no beneficial effect in favor of the neural mobilization. However, patients with lower limb symptoms and signs of nerve root compression appear to benefit from a traction-based treatment intervention[17].

Patients classified in group CS have overriding sensory hypersensitivity generated in part by central pain mechanisms, which is unlikely to respond to manual therapy, as has been shown in patients with whiplash-associated disorders[38]. Patients classified as CS may need to be referred to multimodal treatment programs involving cognitive restructuring and pain medication[39].

Finally, patients in group M with somatic referred leg pain are unlikely to benefit from a treatment approach directed at neural structures in the lumbar spine. These patients would potentially benefit from an intervention derived from a treatment-based classification system[16,18,19].

These examples show that a classification system is not only crucial in order to provide the patients with an explanatory rationale for their disorder but also to select treatments that interact with specific mechanisms[40]. Systems of classification must result in consistent discrimination between groups in order to be valid[41]. The aim of the present study was therefore to investigate the interrater reliability of a new classification system for neural (low back)-related leg pain[27] as the first in a series of studies designed to evaluate the validity of the system.

Methods

This study was approved by Human Research Ethics Committee, Curtin University of Technology, Perth, Western Australia. Data were collected between February 2006 and July 2007. All subjects provided written informed consent.

Subjects

Data for this study were collected from 40 of 77 subjects participating in another study (Figure 2). They were recruited from 162 consecutive patients with low back-related leg pain referred to the physiotherapy department at a private multidisciplinary pain clinic. Patients were referred from within the pain clinic, two orthopedic private practices, and a neurosurgery private practice in Hamburg, Germany. Of the patients referred, 26 were not interested in participating. Those willing to participate were screened for eligibility according to strict inclusion and exclusion criteria. Inclusion criteria were presence of unilateral leg pain of more than 6 weeks and a pain score on the 11-point numerical rating scale of more than 3. Patients were excluded if they fulfilled one of the following criteria: 1) history of lower quadrant surgery or trauma within the past 6 months, 2) nerve root block within the past 4 weeks, 3) history of neuropathic pathology such as diabetes or polyneuropathies, 4) history of vascular disease

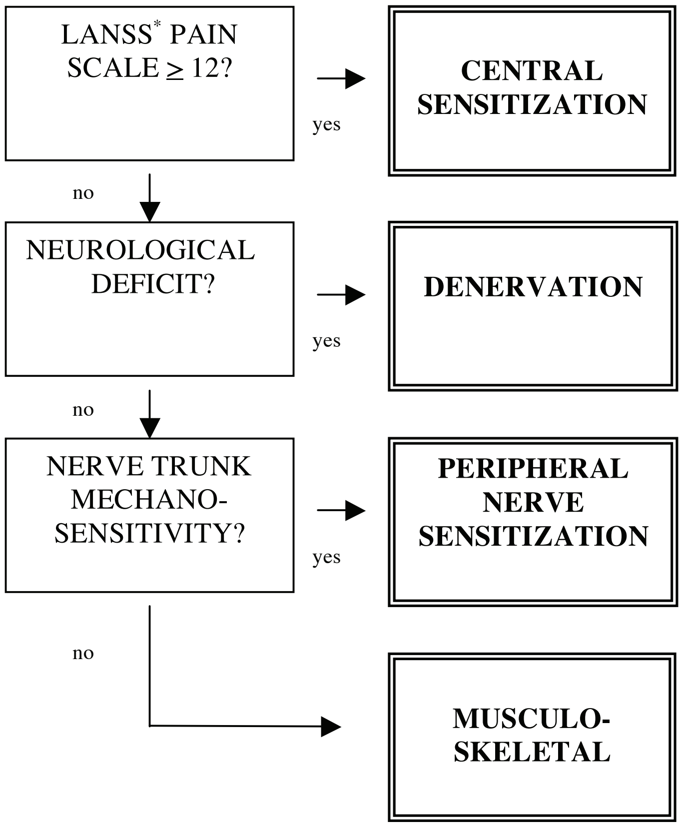

FIGURE 1. Classification algorithm

*LANSS: Leeds Assessment of Neuropathic Symptoms and Signs[42]

in the lower extremities, 5) systemic disease, 6) inflammatory arthropathies, 7) contraindications to manual therapy techniques, and 8) inability to understand written/spoken German. Forty-seven subjects did not meet the selection criteria. Subjective demographic details are shown in Table 1 and data from the physical examination in Table 2.

Examiners

Six experienced physiotherapists who had a mean (SD) of 10 (2.4) years experience in treating musculoskeletal disorders were the examiners for the study. All had undertaken postgraduate education in manual therapy. To standardize the examination procedure, all examiners attended two group and one individual training sessions prior to the start of the study. Additionally, each examiner was provided with a booklet with photos and a description of each of the tests.

Procedure

Each subject was examined on two occasions, once by AS and once by one of the five other participating physiotherapists. Of the five physiotherapists, one examined 12 subjects, two examined 10 subjects, and two examined 4 subjects. Order of testing varied but could not be randomized due to organizational constraints. All subjects were re-examined within 4 days with 70% of the subjects examined twice within a time frame of 48 hours. Each physiotherapist conducted a complete assessment and recorded the findings in a private treatment room. No information regarding the examination or classification of subjects was exchanged between the examiners. Patients were not informed about findings of the first examination to avoid biasing the second examination.

The examination protocol commenced with a subjective examination including questions relating to area of pain, duration of symptoms. and the questionnaire component of the Leeds Assessment for Neuropathic Symptoms and Signs (LANSS)[42]. The physical examination incorporated a) a standard neurological examination screening for motor and sensory deficits; b) the objective part of the LANSS assessment, to examine for allodynia and hyperalgesia[42]; and c) tests for peripheral nerve sensitization[43]. The examination took 30 minutes per subject, including documentation of test results. The individual tests are described in Appendix A. Patients with a LANSS scale score of 12 or higher were classified as CS. Subjects with LANSS score of less than 12 and evidence of frank nerve root damage leading to negative symptoms and signs such as hypoesthesia or failure of motor units were classified as D. Where the first two categories did not apply *and* there were signs and symptoms indicative of increased peripheral nerve mechanosensitivity, patients were classified PNS (Appendix A). Finally, subjects who did not meet the previous criteria were classified as M (Figure 1). Prevalence of classifications was calculated from the results of the first examiner.

Sample Size

The sample size estimates were based on the assumption that 10% of patients would be classified as CS, 10% D, 30% PNS, and 50% M. In this case, to achieve a significant result ($p<0.05$) with a power of above 80%, a sample of 40 subjects measured twice would be sufficient to determine that reliability was at least good ($\kappa > 0.6$)[44,45].

Data Analysis

Data was analyzed using SPSS version 13.0. Percentage agreement, kappa-coefficient (κ), and 95% confidence interval (CI) were calculated to determine the level of agreement between examiners[46].

Results

There was good agreement between examiners (percentage agreement = 80; κ = 0 .72; 95% CI = 0.57 to 0.86). The pro-

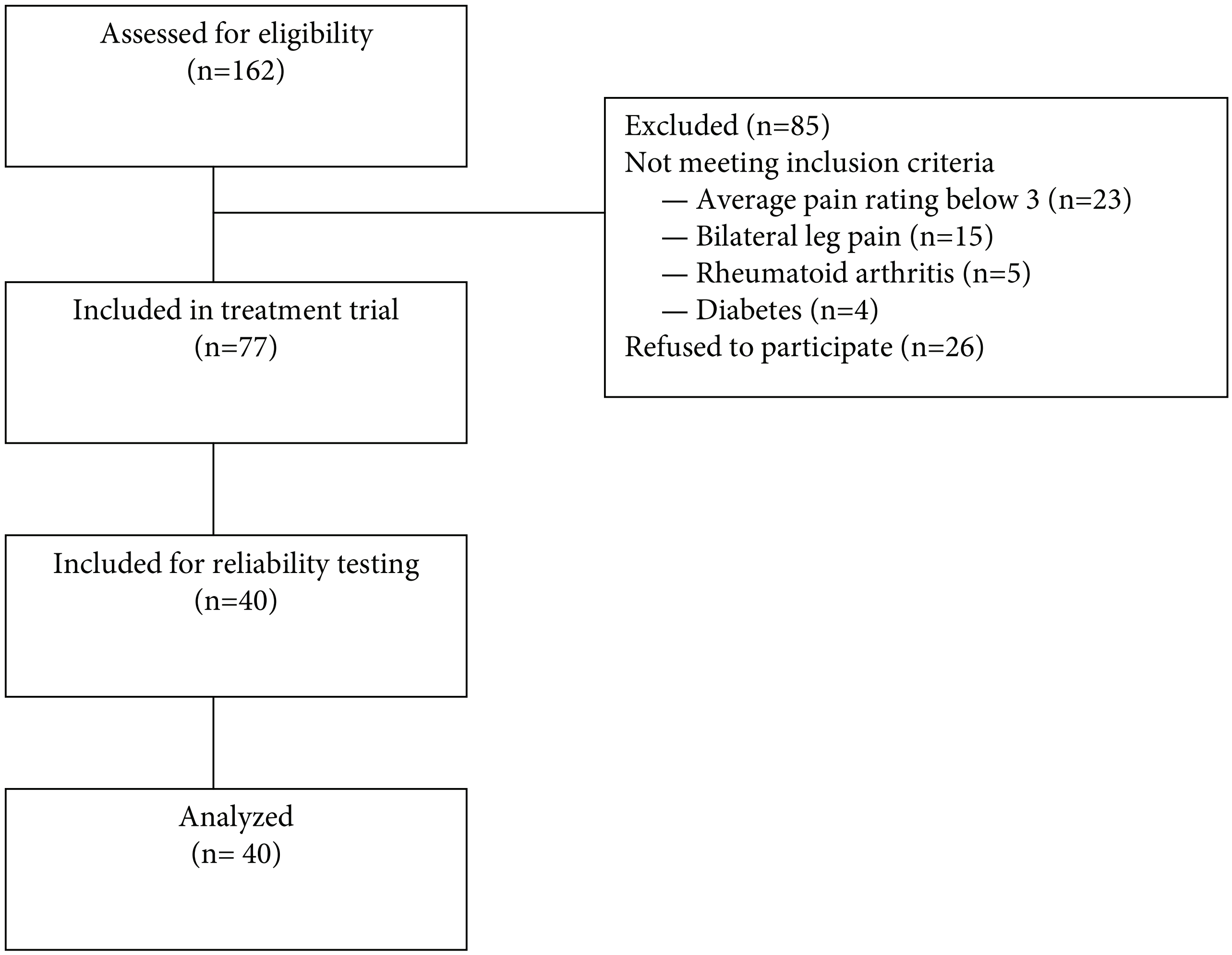

FIGURE 2. Flow diagram

portion of subjects in each classification differed somewhat between examiners as would be expected where agreement was not 100%. Prevalence of the CS, D, and M groups were comparable (27.5%, 32.5%, and 30%, respectively) with a considerably lower prevalence of PNS (10%) (Table 3). Post hoc power analysis showed a power of 88% for detecting a κ of 0.7 with the distribution of groups as described above.

Discussion

Patients with low back-related leg pain can be reliably classified into four groups by experienced musculoskeletal physiotherapists using a new classification system[27] (Figure 1). Reliability is an essential requirement for any valid classification system. Furthermore, to have clinical utility, research results must be generalizable to the clinical environment. Our study was conducted under clinical conditions, with 6 independent examiners assessing consecutive patients with low back-related leg pain. The sample of 40 consecutively recruited subjects (described in Table 1 and 2) can be considered representative of everyday physiotherapy practice in Germany.

Advantages of this classification system are that the classification algorithm relies on routine testing procedures. Experienced clinicians can quickly learn the examination protocol, and the classification algorithm is easy to apply, based on objective decisions for classification into the mutually exclusive categories (for details, see Appendix A). The combination of dichotomous (present/not present) and continuous (e.g., range of motion) measures of outcome minimizes subjective individual interpretations, leading to a clear and easy-to-follow system with good reliability. Thus, clinicians can apply the classification system to their patients without having to learn a new skill-set. Furthermore, they can be confident of the reliability of the classification system when used in the manner described. We believe that the new classification system we are evaluating can be combined with previously reported systems of classifying LBP[13-15]. It was not our intention to replace these, simply to add to them. The purpose of refining these classification systems is to more accurately differentiate neural-related disorders so that

TABLE 1. Subjective characteristics

	Mean (SD)
Number (n)	40
Age	47.8 (13.1)
Gender (% male)	40
Pain below knee (%)	76.3
Pain duration current episode in months (mean/interquartile range)	7.5 (4.0)
Present pain (NRS)	5.0 (1.6)
LANSS score*	8.1 (5.7)
RMDQ	8.0 (4.4)
FABQ	34.7 (19.1)
HADS-A	6.8 (4.1)
HADS-D	5.3 (3.6)

NRS = 11-point numerical rating scale: 0 = no pain to 10 = worst imaginable pain; the score was the mean of three NRS scores: present pain, worst pain, and least pain in the last seven days.
LANSS = Leeds Assessment of Neuropathic Symptoms and Signs
RMDQ = Roland Morris Disability Questionnaire score
FABQ = Fear Avoidance Questionnaire
HADS-A = Hospital Anxiety and Depression Scale—subscale anxiety
HADS-D = Hospital Anxiety and Depression Scale—subscale depression
SD = standard deviation

TABLE 2. Objective characteristics

	CS		*D*		*PNS*		*M*		*Total*	
	n	*%*	*n*	*%*	*n*	*%*	*n*	*%*	*n*	*%*
Reflex deficit	4	33	4	33	0	0	1	8	9	23
Strength deficit	6	50	6	50	0	0	5	41	17	43
PP deficit	8	67	9	75	0	0	4	33	21	53
LT deficit	6	50	7	58	1	25	1	8	15	38
Positive NP	10	83	9	75	4	100	3	27	26	67
Positive SLR	3	25	4	33	4	100	4	36	15	39
Crossed SLR	0	0	1	9	0	0	0	0	1	3
Positive AROMF	5	42	4	36	4	100	0	0	13	33

PP = pinprick; LT = light touch; NP = nerve palpation; SLR = straight leg raise test; AROMF = active range of motion in flexion test; CS = central sensitization; D = denervation; PNS = peripheral nerve sensitization; M = musculoskeletal

TABLE 3. Comparison of findings between physiotherapist examiners

	Second Examiner					
	Central Sensitization	Denervation	Peripheral Nerve Sensitization	Musculo-skeletal	Total	%
First Examiner						
Central Sensitization	10	2	0	0	12	30
Denervation	0	9	0	2	11	27.5
Peripheral Nerve Sensitization	0	1	2	1	4	10
Musculoskeletal	0	2	0	11	13	32.5
Total	10	14	2	14	40	100

treatments such as neural mobilization can be directed to those patients most likely to benefit. Further studies will be required to determine the effectiveness of neural mobilization for patients with PNS.

Despite the simplicity of the classification system, to ensure that it was uniformly executed in the study, all examining therapists participated in pre-study training by TH and AS and were provided with a procedure instruction booklet for reference. Veracity of the data was maintained by strict blinding of the examiners. Other strengths of the study include complete reporting of frequencies of outcome and agreement (Table 3) and data analysis implementing kappa statistics and 95% CIs.

A limitation of our study is that pairing of examiners and the order of testing were not randomized. Instead, the first author was always part of the examiner pair, and order of testing was dictated by organizational circumstances. This limitation may be countered by the fact that this examination process reflects a more typical clinical situation as encountered in everyday practice.

The reliability of a number of examination procedures and five classification systems for patients with nonspecific LBP has been investigated. These studies have recently been reviewed[47], and it was found that results of the five examined LBP classification systems[13-15,48,49] were often conflicting and that interpretation was confounded by variations in the quality and design of the studies. For example, one important point in regards to study design is the procedure of examination. Assessments may be carried out simultaneously or consecutively. In general, for a simultaneous assessment design, both examiners are present during the examination. Alternately, photos, videotapes, and documented subjective examination may be used in place of attendance at the clinical examination. While simultaneous assessments have the advantage of eliminating variations of patients' reports and behavior between two sessions and the effects of repeated testing, which may alter symptom response, these may also introduce bias and thereby artificially inflate the κ value. In contrast, consecutive assessments may be confounded by real changes in the condition of the patients, either secondary to or independent of the initial examination. The interval between consecutive examinations must be individualized to the investigated pathology in the study group and the nature of the examination.

Three of the classification systems reviewed by May et al[47] were relevant in regards to radiating leg pain. The McKenzie classification system[13] was most extensively investigated with four studies providing conflicting evidence on interrater reliability with kappa values ranging from 0.26 to 1.00[50-53].

The patho-anatomic-based diagnostic classification system[15] has shown fair to excellent agreement on syndromes reported by one high-quality study, with kappa values ranging from 0.44 to 1.00 for diagnosis of syndromes but poor reliability for the classification system as a whole (overall agreement on all syndromes for each patient, as syndromes can coexist) with 39% agreement[54]. The treatment-based classification system[14] demonstrated fair to good agreement[55,56] for the overall classification system.

When comparing results of reliability studies, other factors also have to be taken into account. Although the κ coefficient is widely acknowledged as the appropriate statistic for analyzing categorical data, there are some issues that may confound the interpretation of κ values. The κ coefficient indicates the proportion of achieved agreement between examiners beyond agreement by chance alone[46]. Two factors influence the magnitude of the κ coefficient: the prevalence of a symptom or syndrome and the pattern of disagreement (bias) between observers[57]. Feinstein and Cicchetti[57] found two paradoxes in regards to prevalence and bias. The first paradox is the case of either the high or low prevalence of a symptom or syndrome, where agreement per chance would be high and the κ coefficient would be accordingly reduced. A good example for this paradox is found in a study measuring interrater reliability of the McKenzie classification system where both examiners assigned the patients mainly into the derangement syndrome group (90%; 35 of 39 patients)[53]. Although percentage agreement was high at 95%, the κ coefficient was relatively low at 0.60. In our study, the proportion of subjects in the CS, D, and M groups were reasonably evenly distributed (32.5%, 30%, and 27.5%, respectively), whereas prevalence of subjects in the PNS group was low at only 10%. The low prevalence of PNS may have deflated the κ value in our study. The extent to which low prevalence in a study sample is a mirror of the true prevalence in the population or whether it rather reflects the clinicians' diagnostic behavior is questionable[58]. In addition to prevalence, it is also possible that bias may have influenced κ values in our study, but this could not be tested as calculating bias is only suitable for 2x2 tables[58].

Also the number of diagnostic categories should be taken into account when comparing κ statistics for classification systems. The larger the number of categories, the more difficult it is for observers to agree. Additionally, a higher number of examiners will make agreement more complex while improving the quality of a study. Therefore, multiple pairs of examiners are recommended[47,59].

When interpreting κ coefficients as outcomes of reliability research, one also has to consider the condition/pathology for which the investigated test is used. In the case of diagnosing serious pathology, interrater agreement for diagnostic tests should be as good as possible; in this case, even a κ score of above 0.90 may not be satisfactory, as the consequences of a misdiagnosis may be fatal. On the other hand, most measurements in musculoskeletal medicine do not investigate life-threatening conditions; therefore, lower κ values may be considered satisfactory.

Validation of a classification system is a multi-step process[41]. We have established that the classification system we propose is reliable when used by experienced physiotherapists. The next step is to investigate criterion-related validity of the classification system. Once this has been established, studies of the effectiveness of selected treatment techniques can be conducted in more homogenous mechanism-based subgroups of patients.

Conclusion

There was good inter-examiner agreement for the mechanism-based classification system for low back-related leg pain. Therefore, further investigation of the validity of the classification is justified.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgements

The authors would like to thank Dr. Gerd Müller, Kerstin Brass, Lucia Grauel, Frank Kammeier, and Kirsten Keller of the Rückenzentrum am Michel, Hamburg, for their support and assistance in this investigation. In addition, the authors would like to thank Dr. Rebolledo, Dr. Poetsch, and Dr. Thomsen for their support in patient recruitment and Dr. Ritu Gupta for statistical support.

REFERENCES

1. 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;23:1992–1996.
2. Henschke N, Maher CG, Refshauge KM, Das A, McAuley JH. Low back pain research priorities: A survey of primary care practitioners. *BMC Fam Prac* 2007;8:40.
3. Van Boxem K, Van Zundert J, Van Zundert J, Patijn J, van Kleef M. Pseudoradicular and radicular low-back pain: How to diagnose clinically? *Pain* 2008;135:311–312; author reply 313–315.
4. Leffler AS, Hansson P. Letter to the editor of *Pain* on Freynhagen et al: Pseudoradicular and radicular low-back pain: A disease continuum rather than different entities? Answers from quantitative sensory testing. *Pain* 2007;135:65–74.
5. Cavanaugh JM, Weinstein JN. Low back pain: Epidemiology, anatomy and neurophysiology. In Wall PD & Melzack R, eds.

The Textbook of Pain. New York: Churchill Livingstone, 1994.
6. Heliövaara M, Impivaara O, Sievers K. Lumbar disc syndrome in Finland. *J Epid Com Health* 1987;41:251–258.
7. Selim AJ, Ren XS, Fincke G, et al. The importance of radiating leg pain in assessing health outcomes among patients with low back pain: Results from the Veterans Health Study. *Spine* 1998;23:470–474.
8. Freynhagen R, Baron R, Gockel U, Tölle TR. Screening of neuropathic pain components in patients with chronic back pain associated with nerve compression: A prospective observational study (MIPORT). *Curr Med Res Op* 2006;22:529–537.
9. Merskey H, Bogduk N. *Classification of Chronic Pain*. 2nd ed. Seattle, WA: IASP Press, 1994.
10. Bogduk N, McGuirk B. *Medical Management of Acute and Chronic Low Back Pain: An Evidence-Based Approach*. Amsterdam, the Netherlands: Elsevier, 2002.
11. Vroomen PC, de Krom MC, Slofstra PD, Knottnerus JA. Conservative treatment of sciatica: A systematic review. *J Spinal Disord* 2000;13:463–469.
12. Luijsterburg PA, Verhagen AP, Ostelo RW, van Os TA, Peul WC, Koes BW. Effectiveness of conservative treatments for the lumbosacral radicular syndrome: A systematic review. *Eur Spine J* 2007;16:881–899.
13. McKenzie RA. *The Lumbar Spine: Mechanical Diagnosis and Therapy*. Waikanae, NZ: Spinal Publications, 1981.
14. Delitto A, Erhard RE, Bowling RW. A treatment-based classification approach to low back syndrome: Identifying and staging patients for conservative treatment. *Phys Ther* 1995;75:470–485; discussion 485–479.
15. Petersen T, Laslett M, Thorsen H, Manniche C, Ekdahl C, Jacobson S. Diagnostic classification of non-specific low back pain: A new system integrating pathoanatomic and clinical categories. *Physiother Theory Pract* 2003:19;213–237.
16. Long A, Donelson R, Fung T. Does it matter which exercise? A randomized control trial of exercise for low back pain. *Spine* 2004;29:2593–2602.
17. Fritz JM, Lindsay W, Matheson JW, et al. Is there a subgroup of patients with low back pain likely to benefit from mechanical traction? Results of a randomized clinical trial and subgrouping analysis. *Spine* 2007;32: E793–E800.
18. Browder DA, Childs JD, Cleland JA, Fritz JM. Effectiveness of an extension-oriented treatment approach in a subgroup of subjects with low back pain: A randomized clinical trial. *Phys Ther* 2007;87:1608–1618; discussion 1577–1609.
19. Brennan GP, Fritz JM, Hunter SJ, Thackeray A, Delitto A, Erhard RE. Identifying subgroups of patients with acute/subacute "nonspecific" low back pain: Results of a randomized clinical trial. *Spine* 2006;31: 623–631.
20. Campbell JN, Meyer RA. Mechanisms of neuropathic pain. *Neuron* 2006;52:77–92.
21. Balasubramanyan S, Stemkowski PL, Stebbing MJ, Smith PA. Sciatic chronic constriction injury produces cell-type-specific changes in the electrophysiological properties of rat substantia gelatinosa neurons. *J Neurophysiol* 2006;96:579–590.
22. Bove GM, Ransil BJ, Lin HC, Leem JG. Inflammation induces ectopic mechanical sensitivity in axons of nociceptors innervating deep tissues. *J Neurophysiol* 2003; 90:1949–1955.
23. Eliav E, Benoliel R, Tal M. Inflammation with no axonal damage of the rat saphenous nerve trunk induces ectopic discharge and mechanosensitivity in myelinated axons. *Neuroscience Letters* 2001;311:49–52.
24. Eliav E, Herzberg U, Ruda MA, Bennett GJ. Neuropathic pain from an experimental neuritis of the rat sciatic nerve. *Pain* 1999;83:169–182.
25. Woolf CJ, Mannion RJ. Neuropathic pain: Aetiology, symptoms, mechanisms, and management. *Lancet* 1999;353:1959–1964.
26. Woolf CJ. Dissecting out mechanisms responsible for peripheral neuropathic pain: Implications for diagnosis and therapy. *Life Sci* 2004;74:2605–2610.
27. Schäfer A, Hall T, Briffa K. Classification of low back-related leg pain: A proposed patho-mechanism-based approach. *Man Ther* 2009;14:222–230.
28. Elvey RL, Hall TM. Neural tissue evaluation and treatment. In Donatelli R, ed. *Physical Therapy of the Shoulder*. New York: Churchill Livingstone, 1997.
29. Elvey R. Treatment of arm pain associated with abnormal brachial plexus tension. *Austr J Physiother* 1986;32:225–230.
30. Allison GT, Nagy BM, Hall T. A randomized clinical trial of manual therapy for cervico-brachial pain syndrome: A pilot study. *Man Ther* 2002;7:95–102.
31. Coppieters MW, Stappaerts KH, Wouters LL, Janssens K. The immediate effects of a cervical lateral glide treatment technique in patients with neurogenic cervicobrachial pain. *J Orthop Sports Phys Ther* 2003;33: 369–378.
32. Cowell IM, Phillips DR. Effectiveness of manipulative physiotherapy for the treatment of a neurogenic cervicobrachial pain syndrome: A single case study–experimental design. *Man Ther* 2002;7:31–38.
33. Hall T, Elvey R, Davies N, Dutton L, Moog M. Efficacy of manipulative physiotherapy for the treatment of cervicobrachial pain. 73–74. Presentation at 10th Manipulative Physiotherapists Association of Australia Conference. Melbourne, Australia,1997.
34. Cleland JA, Hunt GC, Palmer JA. Effectiveness of neural mobilization in the treatment of a patient with lower extremity neurogenic pain: A single case design. *J Man Manip Ther* 2004;12:143–152.
35. Cleland JA, Childs JD, Palmer JA, Eberhart S. Slump stretching in the management of non-radicular low back pain: A pilot clinical trial. *Man Ther* 2006;11:279–286.
36. Cibulka MT, Aslin K. How to use evidence-based practice to distinguish between three different patients with low back pain. *J Orth Sports Phys Ther* 2001;31:678–688; discussion 689–695.
37. Kitteringham C. The effect of straight leg raise exercises after decompression surgery: A pilot study. *Physiother* 1996;82:115–123.
38. Jull G, Sterling M, Kenardy J, Beller E. Does the presence of sensory hypersensitivity influence outcomes of physical rehabilitation for chronic whiplash? A preliminary RCT. *Pain* 2007;129:28–34.
39. Lidbeck J. Central hyperexcitability in chronic musculoskeletal pain: A conceptual breakthrough with multiple clinical implications. *Pain Res Man* 2002;7:81–92.
40. Woolf CJ, Bennett GJ, Doherty M, et al. Towards a mechanism-based classification of pain. *Pain* 1998;77:227–229.
41. Ford J, Story I, O`Sullivan P, McMeeken J. Classification systems for low back pain: A review of the methodology for development and validation. *Phys Ther Rev* 2007;12:33–42.
42. Bennett M. The LANSS pain scale: The Leeds Assessment of Neuropathic Symptoms and Signs. *Pain* 2001;92:147–157.
43. Hall TM, Elvey RL. Nerve trunk pain: Physical diagnosis and treatment. *Man Ther* 1999;4:63–73.
44. Byrt T. How good is that agreement? *Epidemiology* 1996;7:561.

45. Fleiss JL, Levin BA, Paik MC. *Statistical Methods for Rates and Proportions*. 3rd ed. Wiley Series in Probability and Statistics. Hoboken, NJ: John Wiley & Sons, 2003.
46. Landis JR, Koch GG. The measurement for observer agreement for categorical data. *Biometrics* 1977;33:159–174.
47. May S, Littlewood C, Bishop A. Reliability of procedures used in the physical examination of non-specific low back pain: A systematic review. *Austr J Physiother* 2006; 52:91–102.
48. Sahrmann S. *Diagnosis and Treatment of Movement Impairment Syndromes*. St. Louis, MO: Mosby, 2001.
49. Hall H, McIntosh G, Melles T. A different approach to back pain diagnosis: Identifying a pattern of pain. *Can J Cont Med Ed* 1994;6:31–42.
50. Razmjou H, Kramer JF, Yamada R. Intertester reliability of the McKenzie evaluation in assessing patients with mechanical low-back pain. *J Orthop Sports Phys Ther* 2000; 30:368–383; discussion 384–389.
51. Clare HA, Adams R, Maher CG. Reliability of McKenzie classification of patients with cervical or lumbar pain. *J Manipulative Physiol Ther* 2005;28:122–127.
52. Riddle DL, Rothstein JM. Intertester reliability of McKenzie's classifications of the syndrome types present in patients with low back pain. *Spine* 1993;18:1333–1344.
53. Kilpikoski S, Airaksinen O, Kankaanpaa M, Leminen P, Videman T, Alen M. Interexaminer reliability of low back pain assessment using the McKenzie method. *Spine* 2002; 27:E207–E214.
54. Petersen T, Olsen S, Laslett M, et al. Intertester reliability of a new diagnostic classification system for patients with non-specific low back pain. *Austr J Physiother* 2004;50:85–94.
55. Fritz JM, Brennan GP, Clifford SN, Hunter SJ, Thackeray A. An examination of the reliability of a classification algorithm for subgrouping patients with low back pain. *Spine* 2006;31:77–82.
56. Fritz JM, George S. The use of a classification approach to identify subgroups of patients with acute low back pain: Interrater reliability and short-term treatment outcomes. *Spine* 2000;25:106–114.
57. Feinstein AR, Cicchetti DV. High agreement but low kappa. I. The problems of two paradoxes. *J Clin Epid* 1990;43:543–549.
58. Sim J, Wright CC. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Phys Ther* 2005; 85:257–268.
59. van der Wurff P, Hagmeijer RH, Meyne W. Clinical tests of the sacroiliac joint: A systemic methodological review. Part 1: Reliability. *Man Ther* 2000;5:30–36.
60. Hall TM, Elvey RL. Management of mechanosensitivity of the nervous system in spinal pain syndromes. In Boyling JD & Jull G, eds. *Grieves Modern Manual Therapy*. Edinburgh, UK: Churchill Livingstone, 2004.

APPENDIX A: PHYSICAL EXAMINATION PROCEDURE

Central Sensitization

The subjective examination incorporates the Leeds Assessment of Neuropathic Symptoms and Signs (LANSS) scale, a screening tool for neuropathic symptoms and signs indicative of central processes enhancing sensitivity of the sensory system. Subjects with a score of 12 points or above (of 24) on the LANSS scale were classified in the group Central Sensitization.

Denervation

The physical examination included neurological examination, assessment of active movements, and neural tissue provocation tests. A neurological examination was carried out that included assessment of reflexes, muscular strength, and altered sensitivity for pinprick (Aδ) and light touch (Aβ fibers). These tests were screening for negative signs and were therefore considered positive in the presence of diminished or absent reflexes, muscle power, or sensation on the symptomatic side. Subjects were classified in the Denervation group if they had at least two or more positive tests in two different categories (i.e., reflexes, muscle power, light touch, or pinprick sensitivity).

Peripheral Nerve Sensitization

Following the neurological examination was an investigation of signs indicative of peripheral sensitization of the nerve trunk with enhanced mechano-sensitivity. These signs included restricted range of active movement in standing and hyperalgesia on neural tissue provocation tests (e.g., slump test or SLR, and nerve trunk palpation), both of which had to correlate with the suspected nerve trunk that was mechano-sensitive. Positive neural tissue provocation tests, in the absence of positive symptoms and neurological deficit, were classified as indicative of peripheral sensitization of the nerve trunk with enhanced mechanosensitivity. The flexion active range of motion test (AROMF) was carried out in standing, with the patient's feet slightly apart. The patient was asked to bend forward, while keeping the knees straight, to the first onset of discomfort. Finger tip to floor distance was measured and in case of reproduction of leg pain, the test was repeated with a wedge under the foot of the symptomatic leg to increase ankle dorsiflexion. The AROMF test was considered positive if with increased ankle dorsiflexion reproduction of leg pain was more severe.

The SLR test was carried out in supine lying, with a bubble goniometer (MIE, Leeds, UK) held on the tibial tubercle to measure SLR range of motion. The movement of SLR was carried out to the first onset of discomfort and range of motion recorded. If the patient's leg pain was provoked, then the test was repeated with the ankle held in dorsiflexion. The SLR test was only considered positive if the patient's leg pain was more severe when the SLR test was repeated with ankle dorsiflexion.

The prone knee bend (PKB) test was conducted in prone lying, with the goniometer held in the middle of the tibia to measure knee range of motion. The knee was flexed until the first onset of discomfort. If the patient's leg pain was reproduced, the test was repeated with the patient's spine laterally flexed to the contralateral side. The PKB test was only considered positive if the patient's leg pain was more severe when the PKB test was repeated with the spine in contralateral lateral flexion.

Lower limb peripheral nerve trunks were palpated at the following sites: the sciatic nerve immediately lateral to the ischial tuberosity, the tibial nerve in the mid-line of the popliteal fossa, the common peroneal nerve behind the head of the fibula, and the femoral nerve immediately lateral to the femoral artery level with the inguinal ligament. Nerve palpation tests were considered positive if palpation was more painful on the symptomatic leg compared to the asymptomatic side. All tests were carried out first on the asymptomatic and then on the symptomatic leg. All components had to be positive for Peripheral Nerve Sensitization to be present. The procedure and rationale for this examination protocol are described in more detail elsewhere[60].

AROMF, SLR, or PKB and nerve palpation all had to be positive for the patient to be classified as PNS.