

## Intertester Reliability of a Low Back Pain Classification System

Lynda Wilson, BSc,\* Hamilton Hall, MD, FRCSC,\*†† Greg McIntosh, BHK,\*  
and Tony Melles, BSc, PT\*

**Study Design.** This prospective study of intertester reliability examined pairs of therapists' ability to agree independently on a patient's low back pain diagnosis.

**Objective.** To determine the intertester reliability of a low back pain classification system among experienced and novice clinicians.

**Background.** Many of the disparate categorization schemes for patients with low back pain are purely nominal, assigning designations based on the presumptive source of the problem without providing any practical guide for rehabilitation. A useful classification scheme reliably groups patients into treatment-directing categories.

**Methods.** The study included 204 patients with low back pain referred to 10 clinics across Canada. Paired physiotherapists performed separate physical examinations on each patient. Both examiners then completed a simple ballot choosing one of five pain patterns.

**Results.** Agreement on patient classification by independent examiners was 78.9% ( $\kappa = 0.61$ ).

**Conclusion.** This clinically relevant and clearly defined pain pattern system uses key elements of the history and examination to classify patients with low back pain. The pattern provides a framework for initiating active rehabilitation strategy. Using this approach, clinicians agreed on the categorization of 78.9% of mechanical low back pain cases. [Key words: back pain, classification, examination, intertester reliability, pain patterns] *Spine* 1999;24:248-254

Despite its prevalence and increasing socioeconomic impact, low back pain remains enigmatic. As noted by Sikorski,<sup>42</sup> the management of low back pain has not demonstrably advanced since it became recognized as a problem. Progress has been hampered by the lack of a uniform system for categorizing low back patients into meaningful subgroups. Appropriate patient classification is a necessary preliminary step in defining and managing clinical problems.<sup>27</sup> It is fundamental to the advancement of scientific understanding. Before the validation of a treatment protocol can be completed, the classification system must be in place and verified. Only when patients are classified appropriately can meaningful outcomes measurement commence.

Current classifications of patients with low back pain are based primarily on medical diagnosis. Bernard and

Kirkaldy-Willis<sup>19</sup> contended that it is impossible to formulate a rational therapeutic approach to back pain without concisely diagnosing the pain-producing structures. Yet back pain is a symptom complex that encompasses a spectrum of clinical presentations. Its etiology is multifactorial in nature, with the patients' emotional reaction compounding the difficulty of obtaining a purely anatomic diagnosis. Despite the sophistication of current imaging techniques, the origin of back pain is often conjectural. It is widely acknowledged that a precise anatomic or pathologic etiology can rarely be determined reliably.<sup>3,8-10,23,31,35,38,43,44,46,47</sup>

Insurance systems often oblige clinicians to give a nominal diagnosis when neither a pathologic basis nor an anatomic source of low back pain can be determined. Although incriminating a culpable tissue serves to alleviate the stress of uncertainty for the patient, it fails to generate an appropriate management strategy for maximal patient benefit.

Many disparate methods have emerged to categorize patients with low back pain.<sup>2,3,5-8,10-13,15-17,19-22,24,26-30,33,34,36,38,41-43,45,47-49</sup> A number of these were reviewed by Binkley et al.<sup>4</sup> Most systems fail to provide any specific guide to practical intervention. Consequently, low back patients continue to be regarded as members of an unmanageably heterogeneous group.<sup>40</sup> Cognizant of the dilemma impeding progress in back pain management, Delitto et al<sup>6</sup> asserted that there are few clinical entities for which classification has greater importance than for the ubiquitous problem of low back pain.

Classification schemes that attempt to use anatomic site or pathologic process as a basis for differentiation<sup>3,10,13,19,22,28,29,34,36,41,45,49</sup> are forced to lump a residual 85-90% of patients into one large, nondescript group labeled "nonspecific," "idiopathic" or "other."<sup>1</sup> According to Murtagh,<sup>29</sup> preoccupation with organic causation of symptoms may lead to serious errors in the evaluation and care of patients with back pain.

Effective conservative management depends on a classification that is clinically relevant and sufficiently detailed to provide a logical basis for treatment. Beneficial categories group patients who are homogeneous in response to a specific therapeutic approach. The Quebec Task Force<sup>43</sup> classification scheme reflects the conviction of its expert committee that distinct patterns of reliable clinical findings are the only logical basis for back pain categorization and subsequent management. The McK-

From the \*Canadian Back Institute, Toronto, Ontario, the †Department of Surgery, University of Toronto, and the ‡Orthopaedic and Arthritic Hospital, Toronto, Ontario, Canada.

Acknowledgment date: December 4, 1997.

First revision date: March 13, 1998.

Acceptance date: June 4, 1998.

Device status category: 1.

enzie method of back pain diagnosis classifies patients into three syndromes, giving rise to nine possible categories.<sup>26</sup> Categorization directs treatment decisions. The movements used initially to assess the patient also are the movements used in treatment.

The approach used by the authors is similar to that of McKenzie, but a more inclusive grouping produces a classification consisting of five major patterns of pain. The components included in the pain patterns are chosen to provide the framework for formulating an active management strategy designed to reduce or eliminate pain.

The essential elements of this system are determined by the location of the dominant symptoms and by the particular movements or postures that exacerbate or alleviate the pain. A clinical picture sufficient to initiate therapy is constructed from the information provided by the patient history and physical examination. The history includes a limited number of pertinent questions designed specifically to help identify the precise pattern of pain.<sup>14</sup> A carefully structured physical examination, with an emphasis on repetitive test movements and a focused neurologic examination, screens for serious disease while it supports or refutes the presumptive pain pattern diagnosis.

Two of the authors derived this system on the basis of their combined 35 years of clinical experience with back pain patients. Following a decade of development, the present classification system was implemented in 1992 as the primary tool for determining diagnosis and treatment direction. To ensure content and construct validity, the system was revised biannually for 4 years using clinicians' feedback with more than 100,000 patients before reliability testing was initiated.

Although a patient may possess elements of more than one pattern, it is the dominant pattern that determines initial treatment. As treatment progresses, the dominant pain pattern may shift, and therapy is modified accordingly. The categories are mutually exclusive but not permanent. The first four pain patterns are primarily physical, and the fifth is a pattern of abnormal, pain-focused behavior. A sixth category, "alternative," encompasses the less than 2% of patients with nonmechanical or nonspinal pain. Appendix 1 displays the five patterns of pain.

The system functions by identifying clinical syndromes of back pain. It is not designed nor does it attempt to ascribe anatomic sources to the symptoms. The principal goal of early treatment is pain control and recovery of function. In the absence of relevant pathology such as infection, tumor, or other conditions that produce clinical pictures distinct from the patterns encompassed by this classification, the precise location of the pain generator is of little, or at most secondary, importance. The pattern of pain, not the physical diagnosis, directs treatment. The rapid reduction of pain and corresponding gain in function are both measures of therapeutic success and confirmation of correct pattern selection.

Few back pain classification systems have been examined to determine their interobserver reliability, yet this is an essential criterion. Intertester reliability indicates the ability of different observers to obtain the same answer independently. The basis of any system relies on its ability to produce a consistent result between observers. In 1990, Kilby et al<sup>18</sup> examined the intertester reliability of the McKenzie system of back pain classification and demonstrated a 57% agreement. Riddle and Rothstein<sup>37</sup> pointed out deficiencies in Kilby's study and designed their own study, which yielded a 39% agreement. The authors concluded that the intertester reliability of low back pain assessments using the McKenzie system is poor. This prompted the current reliability study of a less complicated yet equally comprehensive treatment-oriented approach.

The purpose of this study was to determine the intertester reliability of a low back pain classification system among experienced and novice clinicians. The authors hypothesized that this simple, clinically relevant, systematic and clearly defined approach to back pain assessment would minimize interobserver error and result in a high percentage of agreement in patient classification between pairs of therapists.

## ■ Methods

**Examiners.** The examiners for this study were 59 physiotherapists practicing in Canadian Back Institute clinics across Canada. Two distinct groups of examiners were included. One group consisted of 33 skilled and experienced physiotherapists. These therapists had demonstrated proficiency in use of this classification scheme. They had achieved 80% or better on an examination consisting of 10 written case simulations. After this achievement, they exhibited practical competency by performing patient assessments in the presence of this system's principal creators (HH, TM). As a final criterion for selection as an experienced therapist, the candidate's charting of patient information was evaluated as satisfactory by the same authors. This group of physiotherapists had an average of 4.7 years (range, 2.0–9.6 years) of experience with this classification system. The other group of examiners represented 26 new physiotherapists, with an average of 10 months (range, 3–15 months) of experience.

**Sample.** The study consisted of 204 patients referred to 10 Canadian Back Institute clinics across Canada. All subjects who volunteered to take part in the study were required to sign a standard consent and release form. The patients typically present with mechanical spinal pain of musculoskeletal origin. The clinics are primary access rehabilitation facilities that focus on pain control in acute, subacute, and chronic ambulatory populations. Patients with suspected systemic disease and cases sustaining trauma sufficient to produce severe bony injury or major neurologic sequelae are directed elsewhere.

It is advantageous for a comprehensive classification system to have few inclusion and exclusion criteria. The only inclusion criterion was the presence of low back pain with or without associated leg symptoms. Patients with concurrent neck pain were excluded.

The average age of the subjects was  $40.6 \pm 12.5$  years

Table 1. Patient Demographics (n = 204)

Treating Therapist	No. of Subjects	Average Age (yr)	Males (%)	Median Time Since Onset of Pain (days)	Previous Episodes of Back Pain (%)*
Experienced	126	40.8	56.3	24	74.6
Novice	78	40.4	57.7	31	83.1
All	204	40.6	56.9	38	77.8

\* Percentage of patients who reported a history of previous episodes of back pain.

(range, 10–84 years). Males comprised 56.9% of the sample. Overall, this was an acute sample with a few chronic patients skewing the average. Thus, median values for the time from onset of pain to assessment are presented. Previous episodes were reported by 77.8% of patients. Table 1 summarizes the demographic data about the subjects.

**Procedure.** Therapists were paired according to predetermined master schedules prepared by the authors. Three master schedules paired four to six letters of the alphabet in as many ways as possible (Figure 1). Clinicians' names were matched randomly with the letters. Experienced and inexperienced examiner groups were paired independently. On scheduled days, four to six physiotherapists assembled at each site. Paired therapists assessed one patient during a scheduled 65-minute period.

Two weeks before the study day, each of the clinics received an instruction package. The contents consisted of 1) an information page for the reception staff in charge of booking appointments, 2) a schedule detailing time slots and therapist pairings, 3) subject consent and release forms, 4) ballot forms for the therapists to indicate their conclusions, 5) a one-page synopsis of the study methodology, and 6) a comprehensive summary of important points for minimizing bias and ensuring efficiency of the study procedure (Appendix 2).

After the clinics received the package, patients who con-

tacted a participating location for an assessment were informed of the study and asked if they were willing to volunteer. Patients who consented to participate in the investigation were assigned to one of eight scheduled time slots. The staff members who booked the appointments were unaware of the identity of the therapist pairs assigned to each period.

A study coordinator supervised the procedure at every location. At the start of the study day, the coordinator met with the examiners to review the protocol. During the course of the study, the coordinator ensured that the schedule and methodology were followed accurately and efficiently, particularly when participants moved from room to room. The coordinator was particularly concerned with conduct that might bias the study.

For the sake of maintaining internal validity and because the intention of this study was to examine the reliability of this classification in clinical practice, a minimal departure from the usual daily clinic operation was an essential requirement. For temporal and logistic reasons, the patient's clinical history was taken once only, during the first 15 minutes of the assessment. Both examiners, preallocated as assessing therapist (AT) and observing therapist (ObT), recorded the history. The AT interacted with the patient, while the ObT listened and observed. The ObT refrained from interfering in any way with obtaining the patient's history. Both therapists used previously developed standardized assessment forms to record the information independently.

After hearing the complete history, the ObT left the examination room. The study coordinator prohibited subsequent contact between the AT and ObT. The AT completed the patient assessment by performing the structured physical examination. The ObT performed the same physical examination after the AT left the room.

Each therapist independently completed a simple ballot choosing the dominant pattern of pain. Choices included Patterns I through V plus a sixth alternative for patients whose assessment revealed that the pain was either nonmechanical or nonspinal. Ballots were sealed and delivered immediately to the study coordinator. At the end of the eighth period, the study coordinator forwarded all ballots by courier to the authors.

The most common and simple method of measuring and reporting reproducibility is the percentage agreement between two independent observations. The kappa coefficient, which corrects for agreement due to chance, is a more robust reliability statistic. Both methods were used for data analysis. To determine if there was a significant difference in intertester reliability between the experienced and novice examiner groups, an independent samples *t* test was performed. Analysis was performed with the Statistical Package for the Social Sciences software, version 6.0 for Windows (SPSS Inc., Chicago, IL).

	4 THERAPISTS		5 THERAPISTS		6 THERAPISTS		
8:00 TO 8:15	A & B	C & D	A & C	B & D	A & B	C & D	E & F
8:15 TO 8:40	B	D	C	B	B	D	F
8:40 TO 9:05	A	C	A	D	A	C	E
9:05 TO 9:20	A & C	B & D	A & D	B & C	A & C	B & E	D & F
9:20 TO 9:45	A	B	A	C	A	E	F
9:45 TO 10:10	C	D	D	B	C	B	D
10:10 TO 10:25	A & D	B & C	A & B	C & D	A & D	B & F	C & E
10:25 TO 10:50	D	C	A	D	D	B	C
10:50 TO 11:15	A	B	A	C	A	F	E
11:15 TO 11:30	A & B	C & D	E & A	B & D	A & F	B & C	D & E
11:30 TO 11:55	A	C	E	B	A	C	D
11:55 TO 12:20	B	D	A	D	F	B	E
12:20 TO 1:15	L U N C H						
1:15 TO 1:30	A & C	B & D	C & E	A & B	A & E	B & D	C & F
1:30 TO 1:55	C	D	C	A	E	B	F
1:55 TO 2:20	A	B	E	B	A	D	C
2:20 TO 2:35	A & D	B & C	D & E	A & C	A & B	C & D	E & F
2:35 TO 3:00	A	C	D	A	A	C	E
3:00 TO 3:25	D	B	E	C	B	D	F
3:25 TO 3:40	A & B	C & D	E & B	C & D	A & C	B & E	D & F
3:40 TO 4:05	B	D	E	D	C	B	D
4:05 TO 4:30	A	C	B	C	A	E	F
4:30 TO 4:45	A & C	B & D	A & D	B & C	A & D	B & F	C & E
4:45 TO 5:10	A	B	D	B	A	F	E
5:10 TO 5:35	C	D	A	C	D	B	C

Figure 1. Master schedules: therapist pairings.

**Table 2. Frequency of Pain Pattern Classifications**

Therapist 1	Therapist 2						Total
	Pattern 1	Pattern 2	Pattern 3	Pattern 4	Pattern 5	Alternative	
Pattern 1	112	13	5		2	2	134
Pattern 2	10	21			1		32
Pattern 3	5		18				23
Pattern 4				2			2
Pattern 5	2	2			7		11
Alternative	1					1	2
Total	130	36	23	2	10	3	204

## ■ Results

Using 59 examiners and 204 subjects, this pain pattern identification system demonstrated a kappa coefficient of 0.608 ( $P < 0.0001$ ) and an overall agreement of 78.9%. The physiotherapists' experience level did not significantly affect reliability measures. Among the experienced therapists, the kappa coefficient was 0.614 ( $P < 0.0001$ ) and there was 80.2% agreement. For the novice group, the kappa coefficient was 0.60 ( $P < 0.0001$ ) and there was 76.9% agreement (Table 2).

## ■ Discussion

Intertester reliability is a fundamental requirement of a classification system. The usefulness of any classification is vastly diminished by inconsistency in interpretation among the various users. Back pain designations based on poorly defined criteria are destined to yield high intertester error and may lead to inappropriate or unfocused treatment. A good classification system minimizes the effects of inherent clinician biases. An important goal is to optimize reliability without compromising validity.

The current study confirmed that this pain pattern identification system exhibits good intertester reliability, with an overall agreement rate of 79% and a kappa of 0.61 ( $P < 0.0001$ ), indicating highly significant reproducibility between therapists in categorization of pain patterns. A kappa between 0.4 and 0.75 denotes good reproducibility.<sup>39</sup> The intertester reliability was consistent for two groups of physiotherapists significantly dichotomous in their level of experience.

Some bias and inaccuracy is inevitable in any classification system due to the inherent idiosyncrasies of the examiners. A study by Mathew et al<sup>25</sup> confirmed this axiom. Clinicians in their study who had just read identical information failed to make the same predictions regarding operation findings.

Current classification systems require medical information that often is unreliable, unavailable, or nonexistent.<sup>20</sup> Waddell<sup>48</sup> noted that reproducibility of individual history and physical examination items is frequently less than 70% in clinical practice. In a study designed to test the reproducibility of clinical findings in low back pain, Nelson et al<sup>32</sup> observed that the more choice a clinician has in grading a single observation, the greater will be the

resultant interobserver error. If elements used to differentiate syndromes have low reliability, the reliability of final classification is accordingly low.

Classification of a patient into one of several categories depends on a number of observations. As the number of parameters increases, the likelihood of agreement decreases through a multiplication of observational errors. The reliability of the final diagnosis depends on the reproducibility of each step contributing to the decision. The authors' classification system was simplified to the lowest level compatible with use. This system of pain pattern identification emphasizes a small number of reliable, clearly defined, key components of the history and physical examination that focus the therapist's attention on the elements necessary for clinically relevant categorization of a patient with low back pain. Location of dominant pain, the constant or intermittent nature of the symptoms, and aggravating/relieving factors direct treatment. Simplification beyond this point would lead to a higher rate of agreement, but offer no advantage to the larger issue of treatment and management.

The ultimate merit of a classification scheme depends on its potential therapeutic usefulness. A good system maintains relevance while optimizing reliability. Thus, this system of classifying patients by pain pattern achieves a fundamental balance between simplicity and utility, limiting the number of clinical categories to the minimum required to achieve practical value.

There is no need to make explicit reference to specific underlying pathology during the initial assessment and early treatment. The response to therapy is diagnostic. The identification of a particular pain pattern is reinforced by the patient's predictable response to appropriate treatment. The early response to pain control serves as part of the overall diagnostic scheme. Failure to respond in a timely fashion alerts the clinician to the possibility of comorbid conditions or serious underlying pathology.

The reliability of the 28-category classification system of activity-related spinal disorders by the Quebec Task Force on Spinal Disorders<sup>43</sup> has never been tested, but high intertester variation is a probable consequence of this more elaborate categorization. Delitto et al<sup>7</sup> concluded that with adequate training, physical therapists

could reliably place patients into just three basic categories: extension principle, flexion principle, and mobilization principle.

McKenzie's<sup>26</sup> classification method produced an interobserver agreement of 39%. One difference between the current system and the McKenzie approach is the extent to which patients are subdivided into separate categories. The classification of lordosis and detection of lateral shift form an integral part of the McKenzie examination, yet these investigations have been demonstrated to have high interobserver error.<sup>18,32</sup>

The current classification scheme is systematic and well defined, with a minimal requirement for subjective decision making. It appears that these qualities enhance interobserver reliability without diluting clinical quality. Although the system was developed with senior staff, the authors tested a novice group to determine the degree of knowledge, sophistication, and training necessary to employ this classification. The findings indicate that with minimal training, physiotherapists can dependently use this pain pattern identification system to recognize one of five typical patterns.

Reliability without validity is not particularly useful. In view of the heterogeneous nature of back pain and its widely divergent natural history, establishing the value of any specific therapy requires large numbers of patients with consistent and verifiable diagnoses. No single site or practice has the statistical power to validate scientifically a particular management strategy. Through its initial iteration and subsequent extensive modifications, the utility of this classification has been tested empirically as the basis of a treatment algorithm for more than 250,00 patients attending more than 50 different clinic locations. With proof of its reliability, a formal test of its usefulness is possible.

The basis of the current system is the identification of typical presentation patterns. It assumes that if common patterns are recognized and their response to appropriate treatment is predictable, then further investigation is not mandatory. This is a deliberate reversal of the conventional approach to back pain diagnosis in which uncommon but progressive, systemic, or life-threatening conditions must be considered and eliminated before attention is paid to more benign presentations. Most back pain is a mechanical problem that does not reflect serious pathology. A clinical picture that fails to fit one of the five basic patterns is recognized immediately and triggers the necessary concern.

## ■ Conclusion

The current system simplifies the multidimensionality of pain and the complex pathophysiologic mechanisms that contribute to the perception and report of pain to a level that is meaningful to both the clinician and patient. On the basis of this study, the back pain pattern identification system demonstrates a high degree of intertester reliability. This approach to back pain diagnosis allowed clinicians to agree on a categorization of mechanical

back pain in 78.9% of the cases. Developing the necessary skills requires minimal training. This reliability and ease of application provides a foundation for pending outcome studies to determine the prescriptive validity of this system.

## Acknowledgments

The authors thank the staff at CBI Brampton, Hamilton, Kitchener, London, Ottawa, Scarborough, Sudbury, Toronto, Calgary, and Vancouver; and the study coordinators and assessing physiotherapists: Heather Baxter, Michael Bilas, Christina Boyle, Wendy Brand, Janet Brydon, Susan Chang, Nathalie Chartrand, David Clarke, Laura Cook, Myreille Cyr-Hansen, Janet Davis, Kirsten Drew, Cindy Durham, Pam Dyck, Joyce Everett, Natalie Flynn, Sharlene Ford, Brenda Frame, Vanessa Garner, Susan Gibbons, Dominique Gilbert, Annie Girouard, Loreen Gomes, Donna Griffin, Kerry Griffin, Brian Harris, Eric Hoppe, Mark Huschilt, Barbara Hutton, Rosalie Ip, Bob Jacobsen, Karl Jones, Robert Karas, Jill Karch, Jan Kellar, Peter Lanni, Murray Lengyel, Julie Liambotis, Patti Loney, Lisa Luoma-Reddy, Heather MacDonald, Melanie Malisani, Lora Mezzarobba, Jack Miller, Roy Myllari, Monica Perl, Angeles Romero, Steven Savage, Wayne Skitch, Robert Stuart, Brigitte Szybbo, Christopher Szybbo, Sharon Tierney, Helene Ungar, Sheela Veerabhadra, John Ward, Alex Warne, Alison Wirch, and Terri Zacharias.

## References

1. Anderson JAD. Problems of classification of low-back pain. *Rheumatol Rehabil* 1977;16:34-6.
2. Barker ME. A practical classification of spinal pains based on a study of patients seen in British general practice over a five-year period. In: Fairbank JCT, Pynsent PB, eds. *Back Pain: Classification of Syndromes*. New York: Manchester University Press, 1990:23-36.
3. Bernard TN, Kirkaldy-Willis WH. Recognizing specific characteristics of nonspecific low back pain. *Clin Orthop* 1987;217:266-80.
4. Binkley J, Finch E, Hall J, et al. Diagnostic classification of patients with low back pain: Report on a survey of physical therapy experts. *Phys Ther* 1993;73:138-50.
5. Coste J, Paolaggi JB, Spira A. Classification of nonspecific low back pain: II. Clinical diversity of organic forms. *Spine* 1992;17:1038-42.
6. Delitto A, Cibulka MT, Erhard RE, et al. Evidence for use of an extension-mobilization category in acute low back syndrome: A prescriptive validation pilot study. *Phys Ther* 1993;73:216-28.
7. Delitto A, Shulman AD, Rose SJ. On developing expert-based decision-support systems in physical therapy: The NIOSH low back atlas. *Phys Ther* 1989;69:554-8.
8. DeRosa CP, Porterfield JA. A physical therapy model for the treatment of low back pain. *Phys Ther* 1992;72:261-72.
9. Deyo RA. Early diagnostic evaluation of low back pain. *J Gen Intern Med* 1986;1:328-38.
10. Dillane JB, Fry J, Kalton G. Acute back syndrome: A study from general practice. *BMJ* 1966;2:82-4.
11. Fast A. Low back disorders: Conservative management. *Arch Phys Med Rehabil* 1988;69:880-91.
12. Frymoyer JW, Andersson GBJ. Clinical classification. In: *Occupational Low Back Pain*. St. Louis: Mosby-Year Book, 1991.
13. Gardner ADH, Pursell LM, Murty K, Smith DG. The management of the clinical problem of spinal pain with the assistance of a microcomputer. In: Hukins DWL, Mulholland RC, eds. *Back Pain: Methods for Clinical Investigation and Assessment*. New York: Manchester University Press, 1986:23-41.
14. Hall H. A simple approach to back-pain management. *Patient Care* 1992;15:77-91.
15. Hall H, McIntosh G, Melles T. A different approach to back pain diagnosis:

Identifying a pattern of pain. *Can J Continuing Medical Education* 1994;6:31-42.

16. Harper AC, Harper DA, Lambert LJ, et al. Symptoms of impairment, disability, and handicap in low back pain: A taxonomy. *Pain* 1992;50:189-95.

17. Heinrich I, O'Hare H, Sweetman B. Validation aspects of an empirically derived classification for nonspecific low back pain. *Statistician* 1985;34:171-86.

18. Kilby J, Stigniant M, Roberts A. The reliability of back pain assessment by physiotherapists using a "McKenzie algorithm." *Physiotherapy* 1990;76:579-83.

19. Kirkaldy-Willis WH, Hill RJ. A more precise diagnosis for low-back pain. *Spine* 1988;4:102-8.

20. Krause N, Ragland DR. Occupational disability due to low back pain: A new interdisciplinary classification based on a phase model of disability. *Spine* 1994;19:1011-20.

21. Macdonald RS. An osteopathic view of back pain classification. In: Fairbank JCT, Pynsent PB, eds. *Back Pain: Classification of Syndromes*. New York: Manchester University Press, 1990:62-72.

22. Magerl F, Aebi M, Gertzbein SD, et al. A comprehensive classification of thoracic and lumbar injuries. *Eur Spine J* 1994;3:184-201.

23. Margo K. Diagnosis, treatment, and prognosis in patients with low back pain. *Am Fam Physician* 1994;49:171-9.

24. Marras WS, Parnianpour M, Ferguson SA, et al. The Classification of Low Back Disorders via Trunk Motion Assessment. Presented at the International Society for the Study of the Lumbar Spine meeting, June, 1994.

25. Mathew B, Norris D, Hendry D, Waddell G. Artificial intelligence in the diagnosis of low-back pain and sciatica. *Spine* 1987;13:168-72.

26. McKenzie RA. *The Lumbar Spine: Mechanical Diagnosis and Therapy*. Waikanae, New Zealand: Spinal Publications Limited, 1981.

27. Moffroid MT, Haugh LD, Henry SM, Short B. Distinguishable groups of musculoskeletal low back pain patients and asymptomatic control subjects based on physical measures of the NIOSH low back atlas. *Spine* 1994;19:1350-8.

28. Mooney, V. The classification of low back pain. *Ann Med* 1989;21:321-5.

29. Murtagh J. Low back pain: Pitfalls in management. *Aust Fam Physician* 1994;23:604-10.

30. Nachemson A, Andersson GBJ. Classification of low-back pain. *Scand J Work Environ Health* 1982;8:134-6.

31. Nachemson AL. Newest knowledge of low back pain. *Clin Orthop* 1992;279:42-54.

32. Nelson MA, Allen P, Clamp SE, DeDombal FT. Reliability and reproducibility of clinical findings in low-back pain. *Spine* 1979;4:97-101.

33. Nelson MA. The identification of back pain syndromes. In: Hukins DWL, Mulholland RC, eds. *Back Pain: Methods for Clinical Investigation and Assessment*. Manchester, New York: University Press, 1986:13-5.

34. O'Brien JP. Mechanisms of spinal pain. In: Wall PD, Melzack R, eds. *Textbook of Pain*. Edinburgh: Churchill Livingstone, 1984:240-51.

35. Pope MH, Novotny JE. Spinal biomechanics. *J Biomech Eng* 1993;115:569-74.

36. Porter RW. Mechanical disorders of the lumbar spine. *Ann Med* 1989;21:361-6.

37. Riddle DL, Rothstein JM. Inter-tester reliability of McKenzie's classifications of the type of syndrome present in patients with low back pain. *Spine* 1993;18:1333-44.

38. Roland M, Morris R. A study of the natural history of low-back pain: Part II: development of guidelines for trials of treatment in primary care. *Spine* 1983;8:145-50.

39. Rosner B. *Fundamentals of Biostatistics*. 4th ed., Chap. 10. Belmont, CA: Duxbury Press, 1995.

40. Rothstein, JM. Introduction to evidence for use of an extension-mobilization category in acute low back syndrome: A prescriptive validation pilot study. *Phys Ther* 1993;73:216-28.

41. Saunders HD. Classification of musculoskeletal spinal conditions. *J Orthop Sports Phys Ther* 1979;3:3-15.

42. Sikorski JM. A rationalized approach to physiotherapy for low-back pain. *Spine* 1985;10:571-9.

43. Spitzer WO, Leblanc FE, Dupuis M. Scientific approach to the assessment and management of activity-related spinal disorders: A monograph for clinicians. Report of the Quebec Task Force on Spinal Disorders. *Spine* 1987;12:1-59.

44. Spratt KF, Lehmann TR, Weinstein JN, et al. A new approach to the low back physical examination: Behavioral assessment of mechanical signs. *Spine* 1990;15:96-102.

45. Sybert GW. Low back pain disorders. *Trans Assoc Life Insur Med Dir* 1988;71:174-97.

46. Tsang IK. Perspective on low back pain. *Curr Opin Rheumatol* 1993;5:219-23.

47. Vällfors B. Acute, subacute, and chronic low back pain: Clinical symptoms, absenteeism, and working environment. *Scand J Rehabil Med Suppl* 1985;11:1-98.

48. Waddell G, Hamblen DL. The differential diagnosis of backache. *Practitioner* 1983;227:1167-75.

49. World Health Organization. *International Classification of Impairments, Disabilities, and Handicaps: A Manual of Classification Relating to the Consequences of Disease*. Geneva: WHO, 1980.

Address reprint requests to

Greg McIntosh, BHK  
Canadian Back Institute  
1200-330 Front Street West  
Toronto, Ontario M5V 3B7  
Canada  
E-mail: gmcintosh@cbi.ca

## ■ Appendix 1. Five Patterns of Mechanical Back Pain

History	Physical Examination
<p>Pattern 1</p> <p>Back/buttock dominant</p> <p>Worse with flexion</p> <p>Constant or intermittent</p>	<p>Worse with flexion and better in extension (a fast responder)</p> <p>Worse with flexion and worse with extension (a slow responder)</p>
<p>Pattern 2</p> <p>Back/buttock dominant</p> <p>Worse with extension</p> <p>Always intermittent</p> <p>Never worse with flexion (movement or position)</p>	<p>Increase pain with extension</p> <p>No effect or less pain with flexion</p>
<p>Pattern 3</p> <p>Leg dominant—below buttock</p> <p>Leg pain is affected by</p> <p>Back movement/position</p> <p>Always constant</p>	<p>Leg symptoms change with back movement</p> <p>Leg symptoms cannot be abolished with movement or position</p> <p>Leg pain decreases with specific back movement (a fast responder)</p> <p>Leg pain increases with any back movement (a slow responder)</p> <p>Must have a positive irritative test</p>
<p>Pattern 4</p> <p>Leg dominant pain—below buttock</p> <p>Intermittent (short duration)</p> <p>Leg pain increased with activity (walking)</p> <p>Leg pain decreased with posture change</p>	<p>No irritative findings</p>
<p>Pattern 5</p> <p>Manipulative behaviour</p> <p>Pain directed activity</p> <p>Poor sleep pattern</p> <p>Expanding symptom array</p> <p>Constant pain</p> <p>Deterioration instead of anticipated recovery</p> <p>Emotional lability</p> <p>Exaggerated pain response to standard stimuli</p>	<p>Three or more non-organic test groups: confirm only</p> <p>Inability to complete the examination: confirm only</p> <p>Movement discrepancy: confirm-only</p>

## ■ Appendix 2

Distributed Summary of Study Coordinators' Responsibilities

Attention: Study Coordinators

### Practical Guidelines for Careful Supervision of Intertester Reliability Study

Before the start of the first assessment, meet with all participating staff to review the study protocol and the pain pattern classification system.

Ensure that all therapists taking part in the study are aware of the length of the day (8:00 AM to 5:35 PM).

The assessing therapist (AT) interacts with the patient during the history, asking the patient the questions, while the observing therapist (ObT) observes. If, at the end of the history taking, the ObT would like to ask the patient any questions not asked by the AT, he or she may do so at the time of the physical examination. During the history taking, it is important that the ObT not interfere with the assessment. Remind the ObT to avoid body language that might indicate impatience with a history that is taking longer than expected.

During the history taking, the AT may give clues via the line of questioning as to his or her opinion about the pattern. The ObT should be advised to form his or her own opinion and to try not to be biased by the AT.

Ensure that the AT uses the assessment form properly so that the ObT is not handicapped in forming his or her conclusion about the pain pattern.

It is important that no contact take place between the AT and ObT between the two assessments. Body language, facial expressions, and verbal cues by the AT can bias the ObT. Arrange a situation whereby the AT leaves the assessment room in a direction that allows him or her to avoid encountering the ObT. The AT will inform the coordinator that the first assessment has been completed. The coordinator will then inform the ObT that the patient is ready for the second assessment.

Absolutely no discussion is to take place between the AT and ObT until both ballots have been filled out, placed in the envelopes, and delivered to the administrator.

Ensure that each assessment room has any aids (lumber rolls, gowns, etc.) that might be required.

Before leaving the room, the AT must ensure that any props that he or she used are returned to where they originally were found. The patient also should be returned to his or her original position on the edge of the examining table.

Take a copy of any assessment forms wherein the two therapists disagreed about the pattern, and send them to the research department at the head office with the ballots and completed schedule.

## ■ Point of View

William S. Marras, PhD  
Professor and Director  
Biodynamics Laboratory  
The Ohio State University  
Columbus, Ohio

This report represents a much needed form of validity analysis that should be performed for any subjectively based clinical procedure such as the one described here. Assessment of validity is extremely important if a technique is to be adopted as a clinical practice, and this study evaluated one of many forms of validity. The authors assessed the intertester reliability of their own low back pain classification system and found this reliability to be good.

The findings would suggest that the symptom patterns described in Appendix 1 are independent and would result in distinct classifications that might be used for initiating a rehabilitation strategy. The gross statistical analyses might lead one to accept this conclusion. However, one must consider the limitations associated with this study to place these claims in perspective. The patient pool did not represent a balance among all pain patterns. In fact, nearly two thirds of pain patterns

among the patients were associated with pattern 1. Pattern 1 represented the most easily identified pattern and thus biased the results in favor of its intertester reliability. Examination of Table 2 suggests that performance among the other patterns between testers was not nearly as consistent. Thus, the results indicate that only one pain pattern (of those including a significantly large number of patients) had good intertester reliability. A patient population group with a different balance of low back disorder distribution might fare very differently in intertester agreement.

The achievements associated with this effort should be viewed in perspective. The claims of the authors only involve intertester reliability; they do not reflect on the ability of the method to help classify patients into clinically meaningful groups. Only large, independent test set, cross-sectional evaluations and prospective evaluations can address these issues.