Intertester Agreement and Validity of Identifying Lumbar Pain Provocative Movement Patterns Using Active and Passive Accessory Movement Tests

Benjamin Hidalgo, PE, PT, MT, DO,[a] Toby Hall, PT, MT, PhD,[b,c] Henri Nielens, MD, PhD,[d] and Christine Detrembleur, PT, PhD[d]

ABSTRACT

Objective: The purpose of this study was to evaluate the interexaminer agreement and validity of active and passive pain provocation tests in the lumbar spine.

Methods: Two blinded raters examined 36 participants, 18 of whom were asymptomatic and 18 reported subacute nonspecific low back pain (LBP). Two types of pain provocation tests were performed: (1) physiological movements in single (flexion/extension) and, when necessary, combined planes and (2) passive accessory intervertebral movement tests of each lumbar vertebra in prone with the lumbar spine in neutral, flexion, and extension position.

Results: The interobserver agreement in both groups was good to excellent for the identification of flexion ($\kappa = 0.87$-1) or extension ($\kappa = 0.65$-0.74) as the most painful pattern of spinal movement. In healthy participants, 0% was identified as having a flexion provocative pattern and 8.8% were identified as having an extension provocative pattern. In the LBP group, 20% were identified as having a flexion provocative pattern vs 60% with an extension provocative pattern. The average interexaminer agreement for passive accessory intervertebral movement tests in both groups was moderate to excellent ($\kappa = 0.42$-0.83). The examiners showed good sensitivity (0.67-0.87) and specificity (0.82-0.85) to distinguish participants with LBP using this combined examination procedure.

Conclusion: The use of a combination of pain provocative tests was found to have acceptable interexaminer reliability and good validity in identifying the main pain provocative movement pattern and the lumbar segmental level of involvement. These pain provocation tests were able to distinguish participants with LBP from asymptomatic participants and may help clinicians in directing manual therapy treatment. (J Manipulative Physiol Ther 2014;37:105-115)

Key Indexing Terms: *Physical Examination; Diagnosis; Reproducibility of Results; Musculoskeletal Pain; Musculoskeletal Manipulations; Low Back Pain*

Low back pain (LBP) has a high prevalence in Western societies. It is estimated that up to 84% of the European population will experience, at least once in a lifetime, an episode of LBP, with the prevalence of chronic LBP approximately 23%.[1–3]

Most of LBP is described as nonspecific because a radiologically identified cause for pain can only be determined in a small minority of cases. Indeed, there is a poor correlation between findings on radiologic imaging and symptoms, with a radiologic diagnosis identified in only 15% of cases.[2,4,5] Hence, based on imaging, *nonspecific LBP* is defined by the lack of a recognizable, specific pathology and is usually of unknown origin and etiology.[1–3,6] However, despite this evidence, nociceptive factors have a major role in acute and subacute nonspecific LBP conditions. For example, various structures in the lumbar spine are recognized as causative of LBP due to their innervation.[3] In particular the zygapophysial joints, intervertebral disks, and sacroiliac joints have been determined as nociceptive sources in 15%,[7] 40%,[8] and 30% of LBP [2,9,10] cases, respectively. However, the clinical evaluation of patients with LBP should not focus on pathoanatomical data alone.[3] For example, psychosocial factors play a major role in explaining the development of chronic back pain.[3,11] Therefore, generally speaking, in the case of nonspecific LBP, determining a pathoanatomical

diagnosis for pain is unhelpful, particularly when used to drive the management strategy.[12] As a consequence, there has been a call to better define LBP into distinct subgroups by the development of classification systems based on clusters of signs and symptoms relevant to physical therapy.[6,12,13]

Several classification systems for LBP have been proposed, but only 4 systems meet the criteria for tailoring directly manual therapy management and which have been evaluated scientifically.[12,14,15] These 4 systems are the McKenzie (MK) LBP classification system,[16–18] the Treatment-Based Classification system,[6,19,20] the movement system impairment classification for LBP,[14,21] and the motor control impairment or the classification-based cognitive functional approach.[11,22–24] Nevertheless, the best system for subclassification of people with LBP has not yet been determined.

The MK and treatment-based classification systems interpret the patient's symptom behavior with a series of single and repeated spinal movements and sustained postures performed during clinical examination. The goal of the assessment is to identify the directional pattern that worsens and improves the patient's symptoms. These modalities of physical examination provide a basis for the patient's LBP classification and treatment (eg, repeated spinal movements and sustained positions or passive spinal mobilization and manipulation, stabilization exercises, or traction).[14] In all of these classification systems, the sagittal plane is of major importance to determine specific patterns.

Orthopedic manual therapy (OMT) management for an individual patient is driven by evidence-based practice and the results obtained from the clinical examination of the patient together with clinical reasoning.[2,6,25,26] Therefore, the clinical examination should have evidence of sufficient reliability and validity. However, there are few physical assessments that demonstrate evidence of such qualities.[2,4] Furthermore, there is a generally considered poor correlation between movement impairment and the presence and severity of LBP.[11,24]

There are at least 3 general domains in the clinical assessment of articular dysfunction in LBP: observation of movement and posture, motion palpation for spinal segmental mobility, and pain provocation tests.[4] Investigations of the reliability of these procedures indicate greater reliability for tests of pain or symptom provocation rather than observation or motion palpation.[2,4,12,13,27] For example, a systematic review reported moderate evidence regarding the identification of bony landmarks by palpation, and weak evidence for the evaluation of segmental mobility and segmental dysfunction requiring treatment in the lumbar spine.[4]

The literature suggests that those tests that are the most reproducible, in clinical examination of the lumbar spine, are those that are based on symptom repro-

Table 1. *Baseline characteristics of asymptomatic participants and those with nonspecific LBP*

	Healthy asymptomatic participants (n = 18), mean (SD)	Participants with LBP (n = 18), mean (SD)
Height (cm)	170.1 (9.2)	172.7 (6.5)
Weight (kg)	67.2 (11.9)	70.3 (12.3)
BMI (kg/m^2)	23.1 (2.9)	23.5 (3.4)
Age (y)	32 (8.3)	38 (9.8)
Sex (female)	9	9
Present pain (VAS)	–	2.9 (1.3)
Roland Morris Disability score	–	8.4 (3.3)
Duration of pain (wk)	–	6.3 (2.9)

LBP, low back pain; *VAS*, visual analog scale.

duction.[2,4,17,19,27] More specifically, the interexaminer reliability of pain response during repeated lumbar spinal movements (in flexion/extension) is the only procedure to show moderate evidence of high reliability.[4] Therefore, when a physical examination is based on the response to symptoms, reliability is good, whereas when it is based on palpation to detect mobility, reliability is generally low.[4,27] Moreover, several studies of good methodological quality have demonstrated the validity of movement tests to discriminate people with LBP from healthy asymptomatic participants using tests of active spinal movement.[5,22,24,28]

One form of pain provocation testing that is commonly used in OMT clinical examination is active movement tests in single or combined planes. The concept of combined movement (CM) testing was originally developed by Edwards[29] and is an expansion of the routine clinical examination. Another form of pain provocation testing is passive accessory intervertebral movement (PAIVM) testing. In the concept proposed by Edwards, information gained from single and combined plane active movement examination is used together with PAIVM tests performed in different lumbar spine positions to determine a pain provocative direction that is more specific to the patient's problem and is also more functional.[2,25,29–31] This pain provocative direction directs manual therapy management, which aims to reduce pain through restoring pain-free range of motion in the specific direction. There have been no studies to date that have investigated the interexaminer agreement or validity of this approach to examination of LBP.

Therefore, the purpose of this study was to investigate the reliability and validity of pain provocation tests to identify a pain provocative direction. Specifically, we examined the combination of active trunk movements with PAIVM of the lumbar spine to determine a pain provocative direction of flexion and/or extension and the involved lumbar levels. We also sought to investigate whether these tests of pain
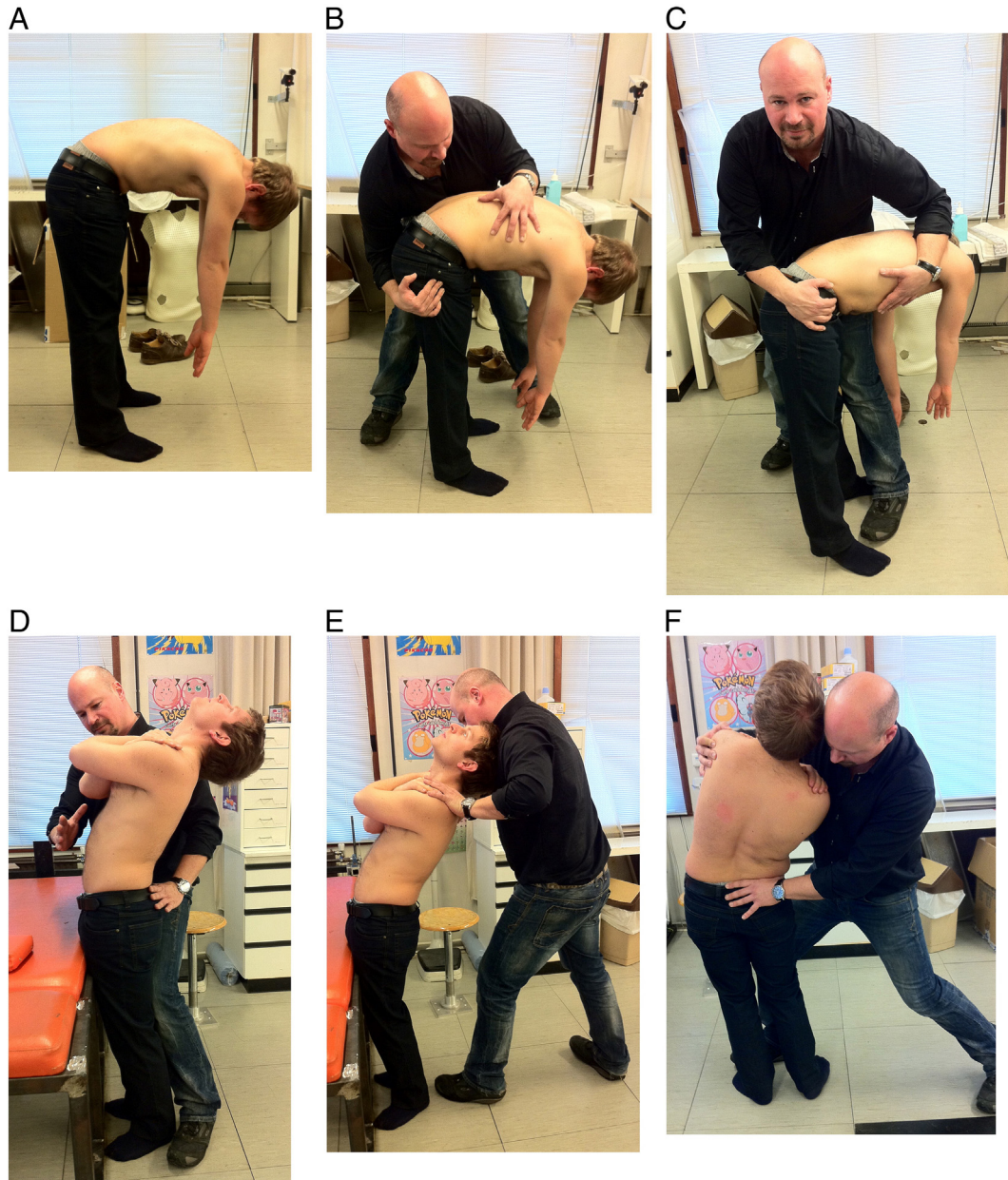
Fig 1. *Single and combined active trunk movement. A, Active flexion with repeated movements (maximum 10 repetitions) and sustained position. B, Sustained flexion with overpressure. C, CMs in primary flexion then secondary lateral flexion left. D, Active extension with repeated movements (maximum 10 repetitions) and sustained position. E, Sustained extension with overpressure. F, CMs in primary extension then secondary lateral flexion right. (Color version of figure is available online.)*

provocation could be used to distinguish participants with LBP from asymptomatic participants.

METHODS

Study Population

Two groups of participants were investigated (Table 1): a group of healthy asymptomatic participants (n = 18) and a group of patients with subacute nonspecific LBP (n = 18).

The inclusion criteria for the asymptomatic group were as follows: aged between 20 and 65 years, body mass index (BMI) less than 30 kg m^{-2} and no back pain for at least 6 months. Asymptomatic participants were recruited on a voluntary basis in response to posters placed around the hospital. The inclusion criteria for the LBP group were as follows: aged between 20 and 65 years, BMI less than 30 kg m^{-2}, and the presence of nonspecific LBP for at least 4 weeks.[32] A medical doctor confirmed the diagnosis of nonspecific LBP. The Roland Morris
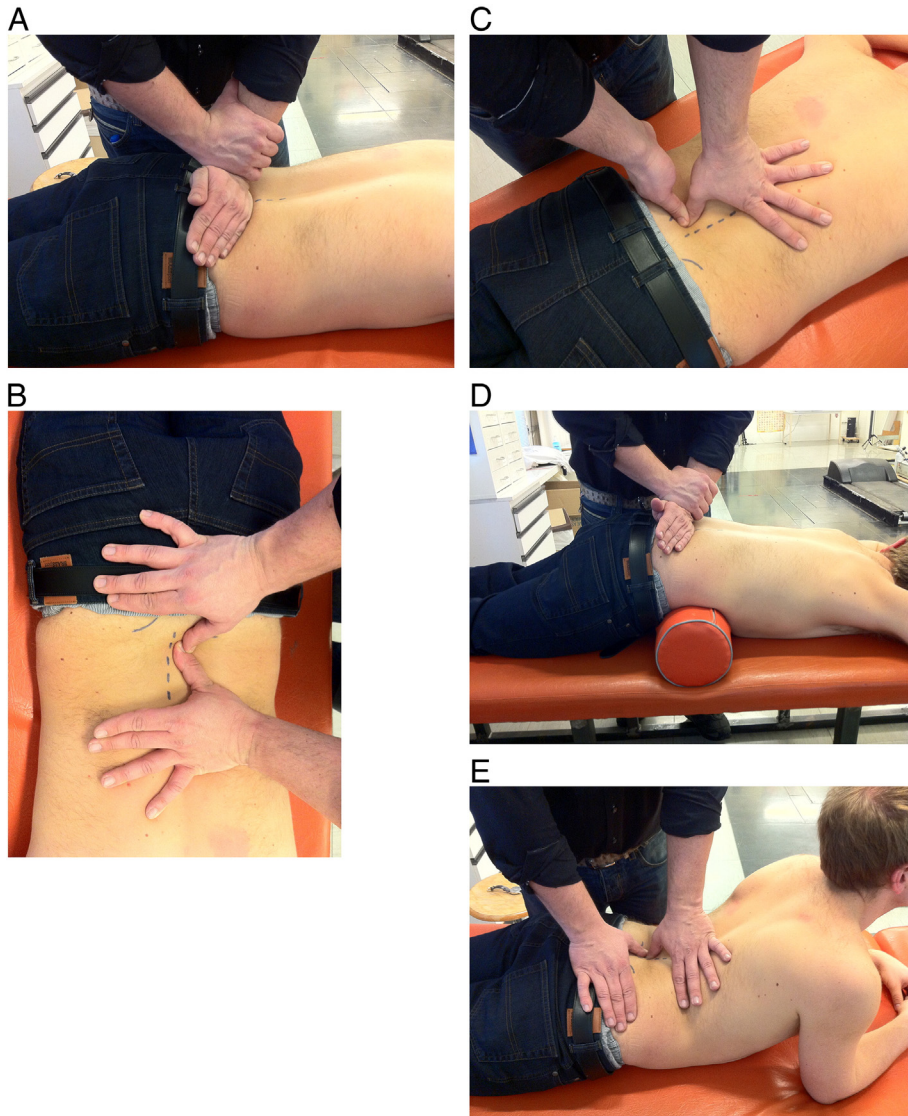
Fig 2. *Passive accessory intervertebral movement tests. A, PA pressure on spinous process in neutral prone position. B, Lateral pressure on the spinous process in neutral prone position. C, PA pressure on the zygoapophyseal joint in neutral position. D, PA pressure on a spinous process in flexion prone position. E, Lateral pressure on the spinous process in extension. (Color version of figure is available online.)*

Disability Questionnaire was used to measure disability. Exclusion criteria included the presence of red flags, rheumatologic diseases, neurologic deficits, and a history of spinal surgery. Patients were recruited from outpatient clinics at the Saint-Luc University Hospital (Belgium) and had symptoms in the low back area and/or irradiation into the lower limb but not below the knee. Patients were excluded from the study if they had a visual analog scale score for pain greater than or equal to 7/10 on the day of the experiment. Each patient participated in the study on a voluntary basis and provided written informed consent, following the principles of the Declaration of Helsinki. This study had ethical approval from the Commission d'éthique Hospitalo-Facultaire de l'Université de Louvain-La-Neuve, Brussels, Belgium.

Assessment Procedures

Two blinded observers with postgraduate qualifications in manual therapy performed the evaluation. One examiner had 10-year postgraduate clinical expertise in OMT, and the other was a novice (degree in OMT but without clinical experience). A third observer recruited the participants, confirmed participant eligibility for each group, and was the only person aware of the participant's group status.

Each participant was examined on a single occasion by both examiners who were blind to each other. During this assessment, 2 types of pain provocation tests were performed: first, active movement pain provocation tests in standing and, second, PAIVM tests in prone. Active movement comprised single-plane repeated active trunk

Table 2. *Reliability of pain provocation direction during trunk movements*

	Healthy asymptomatic participants				Participants with LBP			
	%A	κ	PABAK	%+	%A	κ	PABAK	%+
Repeated flexion	100	–	1	0	86.7	0.45	0.73	13.3
Flexion with overpressure	100	–	1	0	80	0.34	0.6	16.6
CM (F-LFR) [a][b]	100	–	1	0	100	–	1	0
CM (F-LFL) [a][b]	100	–	1	0	93.3	0	0.87	3.3
Repeated extension	100	–	1	0	100	–	1	6.6
Extension with overpressure	94.1	0	0.88	5.8	80	0.65	0.62	50
CM (E-LFR) [a][c]	82.4	−0.09	0.65	2.9	93.3	0	0.87	3.3
CM (E-LFL) [a][c]	94.1	0	0.91	2.9	93.3	0	0.87	3.3
Identification of flexion pattern	100	–	1	0	93.3	0.63	0.87	20
Identification of extension pattern	82.4	0	0.65	8.8	86.7	0.74	0.73	60

%A, percentage agreement between raters; *%+*, percentage of participants with a positive test result; *E-LFL*, extension then lateral flexion left; *E-LFR*, extension then lateral flexion right; *F-LFL*, flexion then lateral flexion left; *F-LFR*, flexion then lateral flexion right; *PABAK*, prevalence-adjusted bias-adjusted κ.

[a] Only if previous test results were negative.

[b] Tests performed in all healthy participants and in 71% of participants with LBP.

[c] Tests performed in 94% of healthy participants and in 43% of participants with LBP.

movements including sustained positioning and overpressure in flexion and extension to reproduce or increase pain. If there were no pain during these spinal movements, then CMs of the trunk were performed.[2,20,25,29,31,33] Passive accessory intervertebral movement tests were applied to each lumbar vertebrae (from L1 to L5) in prone position with the lumbar spine in neutral, flexion, and extension.[27,29,31,33,34] Both examiners performed all tests successively on the same day, with a short break for the participant between examiners (10 minutes). The following 2 factors were randomized: order of examiner, and the order of PAIVM tests or active movement tests.

Battery of Tests

Participants were asked not to inform the examiner of their group allocation. In addition, examiners used standardized communication to question participants in a similar manner on the presence of pain.

SINGLE AND COMBINED ACTIVE TRUNK MOVEMENT

The aim of the active movement examination protocol was to provoke pain (and therefore identify the pain provocative direction), in such a way that progressive strains were placed on the lumbar spine as follows: the patient started in a standing position and performed a maximum of 10 trunk flexion movements, maintaining the knees in extension (Fig 1A). The participant determined a comfortable movement velocity. During the repetition until the 10th, the examiner asked the participant about the presence of pain onset or pain increase. If no pain was provoked, then the participant's spine was sustained in an end-range flexion position for a maximum of 10 seconds, or until pain was provoked.[2,16,17,19,33] In the absence of pain onset or pain increase, the examiner then applied overpressure (Fig 1B) and again seeking the status of pain. The same

procedure was repeated for the movement of trunk extension (Fig 1D, E). If these procedures did not influence pain, then CMs were evaluated.[29,31,33,35] The participant was directed to move to a position of flexion or extension, combined with assisted lateral flexion to the left and right (Fig 1C, F). The CMs used were as follows: flexion with lateral flexion left, flexion with lateral flexion right, extension with lateral flexion left, and extension with lateral flexion right.[29,33]

Following these routine clinical examination procedures, the examiners were required to identify the most pain provocative direction (flexion or extension) for each participant.

Passive Accessory Intervertebral Movement

Prior to the examination, the skin overlying each lumbar spinous process was marked with a visible dermographic pencil using a previously developed method.[27] After this, 5 types of oscillatory PAIVM were applied to each lumbar vertebra: posteroanterior (PA) pressure on each spinous process (Fig 2A), as well as PA pressure on the left and right zygoapophyseal joints (Fig 2B) and lateral pressures applied on the left and right sides of each spinous process (Fig 2C). The method of application has been previously described.[16,20,27,31,33,34] All accessory motion tests were applied in 3 different prone positions: neutral (eg, Fig 2A), flexed over a 20-cm cushion cylinder (eg, Fig 2D), and an extended position achieved through the patient resting on their elbows (eg, Fig 2E).[29,31]

To improve the standardization of force applied by the examiners,[36] each PAIVM test was standardized according to the grades of Maitland.[33,34] The grades applied were progressive oscillatory pressure from grades III to IV. The end point for each test was either pain or end-range resistance with a grade IV pressure, whichever came first. Although these tests are used to assess for hypomobility or

Table 3. *Reliability of pain provocation with PAIVM from L1 to L5*

Healthy asymptomatic group

	PA on spinous process						PA on zygapophyseal joint right						PA on zygapophyseal joint left		
	% A	CI95%	κ	CI95%	PABAK	CI95%	% A	CI95%	κ	CI95%	PABAK	CI95%	% A	CI95%	κ
μ	87.9	85.9-89.9	0.21	0.10-0.32	0.76	0.72-0.80	85.9	83.2-88.5	0.29	0.15-0.43	0.72	0.67-0.77	87.9	87.1-88.6	0.33

LBP group

	PA on spinous process						PA on zygapophyseal joint right						PA on zygapophyseal joint left		
	% A	CI95%	κ	CI95%	PABAK	CI95%	% A	CI95%	κ	CI95%	PABAK	CI95%	% A	CI95%	κ
μ	72.5	65.7-79.3	0.43	0.31-0.54	0.45	0.31-0.58	73.8	71.5-76.1	0.42	0.34-0.49	0.48	0.43-0.52	73.4	70.8-76.0	0.43

%A, percentage agreement between raters; *CI*, confidence interval; *LBP*, low back pain; *PA*, posteroanterior; *PABAK*, prevalence-adjusted bias-adjusted κ; μ, mean of neutral, flexion, and extension prone positions (from L1 to L5).

hypermobility, this aspect of testing was not included and we only assessed for pain provocation. Any pain response was recorded as a positive response. The examiner recorded the dichotomous pain response (present or not), vertebral level at which pain was provoked, and type of accessory movement that provoked pain. The manual examination by PAIVM tests when accompanied by a verbal participant response had previously been demonstrated to be highly accurate in detecting the lumbar segmental level responsible for a participant complaint.[37,38] All vertebrae were tested from L1 to L5, and all accessory movements performed on each vertebral level.

CLINICAL CLASSIFICATION RULE

Both examining therapists were required to state whether the participant they had tested had LBP or not, determined by the presence of a painful pattern of flexion or extension coupled with pain on PAIVM tests.

Hence, a clinical classification rule (CCR) was developed to identify the presence of LBP. This consisted of 3 criteria that were all required to be positive:

Criteria 1: active movement tests. A predominant pain provocative movement direction (flexion or extension) during single, repeated, sustained, or overpressure tests (Fig 1A, B, D, E), or if required in a CM direction (flexion or extension combined then with lateral flexion right or left; Fig 1C, F). After these single or CM tests, the assessors had to establish the most painful pattern of spinal movement: that is, positive = flexion or extension and negative = no painful pattern.
Criteria 2: passive movement tests. At least 2 adjacent vertebral levels provoked pain on PAIVM tests (Fig 2A-C): that is, positive = 2 painful adjacent vertebral levels and negative = 0 or 1 painful vertebral level.
Criteria 3: pain provoked by PAIVM was made worse at the specific vertebral level, by flexing or extending the spine (Fig 2D, E), with the direction in concordance with

the direction of active pain provocative movement previously identified in criteria 1: that is, positive = concordance and negative = no concordance.

STATISTICAL ANALYSIS

Interexaminer agreement for single and CMs and the identification of the pain provocative direction (Table 2) as well as PAIVM tests (Table 3) were calculated by using the percentage of agreement (%A) and κ test (MedCalc software, version 11.5; MedCalc, Mariakerke, Belgium). In some situations, when the prevalence of a given response to a test is either very high or very low, the interpretation of the κ statistic does not satisfactorily reflect the true level of agreement.[39] Other statistical tools have been developed to account for this, such as "prevalence-adjusted bias-adjusted κ" (PABAK),[40] which corrects for this type of bias. κ and PABAK were interpreted according to the classification of Blum et al[41] (Table 4).

The validity of our CCR for each examiner was determined by assessing sensitivity and specificity using the following equations:

$$Sensitivity = true\ positives/(true\ positives + false\ negatives)$$

$$Specificity = true\ negatives/(true\ negatives + false\ positives)$$

Sensitivity refers to the ability of the CCR to correctly identify those patients with LBP. *Specificity* refers to the ability of the CCR to correctly identify those patients without LBP.[37]

One-way analysis of variance (ANOVA) with all pairwise multiple comparison procedures (Holm-Sidak method) (Sigmastat 3.5; Systat Software, Inc, San Jose, CA) was performed in each group on the prevalence of positive responses during PAIVM with comparison for factor 2: levels of vertebra (L1, L2, L3, L4, L5). Two-way ANOVA with all pairwise comparison procedures (Holm-Sidak method) was performed as well with comparison for

Table 3. (*continued*)

| CI95% | PABAK | CI95% | Lateral pressure right side of spinous process | | | | | | Lateral pressure left side of spinous process | | | | | |
			% A	CI95%	κ	CI95%	PABAK	CI95%	% A	CI95%	κ	CI95%	PABAK	CI95%
0.14-0.52	0.75	0.74-0.77	91.4	90.6-92.2	0.45	0.33-0.57	0.83	0.81-0.85	91.4	89.3-93.3	0.42	0.39-0.46	0.83	0.79-0.87

| CI95% | PABAK | CI95% | Lateral pressure right side of spinous process | | | | | | Lateral pressure left side of spinous process | | | | | |
			% A	CI95%	κ	CI95%	PABAK	CI95%	% A	CI95%	κ	CI95%	PABAK	CI95%
0.36-0.50	0.46	0.41-0.52	77.1	74.1-80.2	0.43	0.33-0.52	0.45	0.32-0.58	76.0	68.1-84.0	0.46	0.29-0.63	0.52	0.36-0.68

factor 1: groups (Healthy and LBP) and for factor 2: levels of vertebra (Table 5).

All statistical analyses were performed by the first author.

RESULTS

Interobserver Agreement

Single and Combined Active Trunk Movement Tests in Standing. In healthy participants, the interexaminer agreement of classification of flexion or extension pattern was good to excellent, with PABAK values of 0.65 to 1.00 and %A between examiners ranging from 82.4% to 100%. Moreover, 0% of healthy participants were identified as having a flexion pattern, whereas 8.8% were identified as having an extension pattern. In participants with LBP, the interexaminer agreement of classification of flexion or extension pattern was good, with PABAK values of 0.73 to 0.87 and %A ranging from 86.7% to 93.3%. In the LBP group, 20% were identified as having a flexion pattern, whereas 60% had an extension pattern (Table 2).

Passive Accessory Intervertebral Movements. Interexaminer agreement for PAIVM from L1 to L5 in healthy participants (neutral, flexion, and extension positions) was good for PA pressure on the spinous process, with PABAK of 0.76 and a %A of 87.9%; for PA pressure on the right zygoapophyseal joint, agreement was good, with PABAK of 0.72 and 85.9%A; for PA pressure on left zygoapophyseal joint, agreement was good, with PABAK of 0.75 and 87.9%A; and for lateral pressure on the right and left sides of the spinous process, agreement was excellent, with PABAK of 0.83 and 91.4%A, respectively (Table 3).

The interexaminer agreement for PAIVM from L1 to L5 in participants with LBP (neutral, flexion, and extension positions) for PA pressure on the spinous process was moderate, with PABAK of 0.45 and 72.5%A; for PA pressure on the right zygoapophyseal joint, agreement was moderate, with PABAK of 0.48 and 73.8%A; for PA pressure on left zygoapophyseal joint, agreement was

moderate with PABAK of 0.46 and 73.4%A; for lateral pressure on the right side of the spinous process, agreement was moderate, with PABAK of 0.45 and 77.1%A; and for lateral pressure on the left sides of the spinous process, agreement was moderate, with PABAK of 0.52 and 76.0%A.

One-way ANOVA in the asymptomatic group showed that there was a significant difference ($P < .001$) between levels of vertebra concerning the prevalence of positive tests and that L5 was significantly different from L1, L2, L3, and L4. Similar results ($P < .001$) in the LBP group were found, with significant differences between L5 and L1, L2; L4 and L1; and L3 and L1. Two-way ANOVA determined that the prevalence of positive tests was different ($P < .001$) between both groups (Table 5).

The detailed values of PAIVM, interexaminer agreement, and prevalence of positive tests by vertebra are respectively presented in Tables 3 and 5.

Diagnostic Accuracy of LBP Classification

The sensitivity and specificity of identifying a person with LBP using the proposed CCR were, respectively, 0.87 and 0.82 for the experienced examiner and 0.67 and 0.85 for the novice. The combined sensitivity and specificity were, respectively, 0.77 and 0.84.

DISCUSSION

In accordance with previous studies,[2,4,5,17,19,22,23] this study found good interexaminer agreement for active movement tests of the trunk in asymptomatic participants and patients with LBP. The reliability of identification of the most painful pattern (flexion or extension) was also found to be at least good, which is comparable with other studies investigating movement classification systems in people with LBP.[5,17,19,20,22,23]

In the present study, CM tests were applied in a higher percentage of asymptomatic participants than in patients

Table 4. *Interpretation of the κ according to the classification of Blum et al[41]*

Agreement	κ
Excellent	0.81-1
Good	0.61-0.80
Moderate	0.41-0.60
Weak	0.21-0.40
Negligible	0-0.20
Poor	<0

with LBP. In our protocol, we decided to stop the progressive strain of spinal movement at pain onset or pain increase. Most participants with LBP experienced pain during single, repeated, or sustained active movement tests or with the application of overpressure. Hence, CM tests were less frequently required in participants with LBP. The reproducibility of CM testing has been studied by Haswell et al.[35] Paradoxically, the results showed poor agreement between raters, which is in contrast to our results. The reason for this difference remains unclear and may be related to differences in the frequency that CM tests were performed or alternatively in differences in examiners training and participants characteristics.

In accordance with a previous study,[42] the prevalence during clinical examination of positive responses to PAIVM tests was higher for the lower lumbar vertebrae than for the upper lumbar vertebrae in both groups. However, the results of a 2-way ANOVA demonstrated that the prevalence of positive responses from testing each lumbar vertebra was more frequent in participants with LBP, probably because of specific underlying pain mechanisms in this population. The topographical differences in pain responses in both groups could perhaps be explained by the increased predominance of biomechanical strain and pathoanatomical features in the lower lumbar area when compared with upper lumbar levels.[43,44] Another explanation might be due to the transition from the mobile lumbar spine to relatively rigid pelvis, placing physiological stress on the lower lumbar segments, sensitizing those segments. The presence of positive PAIVM tests in asymptomatic people highlights the importance of identifying more than 1 vertebral level as symptomatic, as adopted in this study. In addition, this highlights the importance of using a combination of factors such as the CCR when distinguishing people with LBP.

There are few studies, to our knowledge, that have reported on the relative frequency to which each vertebral level contributes to LBP. However, based on the clinical observation of vertebral levels that commonly receive surgery, zygoapophyseal injections, or intervertebral discography, the lower lumbar spine would appear to be the more common source of symptoms.[42] The precise location of pain origin in the spine of patients with LBP is of major importance to manual therapists. The L4 and L5 vertebral levels, which are frequently found to have pathology in

LBP people, have also been reported to be the most common vertebral levels to provoke concordant pain during epiduroscopy.[42]

The results for interexaminer agreement for tests of PAIVM are at least comparable with or better than those obtained in previous studies.[27,36–38,45] Previous studies reported either weak to moderate or moderate to good agreement for intraobserver reliability when testing PA pressures only on the spinous process ("spring test"). Schneider et al[27] also reported similar levels of interexaminer agreement for palpation of the zygoapophyseal joints, although they grouped levels into upper and lower lumbar levels rather than a specific vertebral level.

Similarly to Phillips and Twomey[37] and in terms of manual diagnostic accuracy, the validity (sensitivity and specificity) of our CCR to identify people with LBP can be rated as good, especially for the evaluator with better clinical experience. The sensitivity for the novice examiner was slightly less than that found for the experienced examiner. Therefore, the novice examiner should be more careful to interpret a negative result (because of a higher chance of false-negative findings) when testing LBP people with the CCR. The results of this study suggest that pain provocation using the CCR, when accompanied by a verbal participant response,[37] is an important component of physical examination to distinguish people with LBP from healthy participants. This information may be important when one considers the poor correlation between medical imaging and LBP, and the poor correlation between information gained from magnetic resonance imaging and clinical examination in people who have LBP.[5,42,46]

Limitations and Future Studies

There is a major limitation of our report. Validity is usually determined by comparing the results of a new test against a criterion standard. Unfortunately, there is no pathoanatomical criterion standard in the case of nonspecific LBP.[3] Furthermore, in the present study, the patient's own report of pain was used during pain provocation tests because verbal response during manual diagnosis has previously been validated in unilevel lumbar spinal block procedures.[37] This approach may constitute an important limitation of internal validity, but because the method is easily transferable to a clinical context, it provides substantial external application. Nevertheless, a recent study highlights the interest of epiduroscopy as an external reference to help diagnose the vertebral level of pain in people with LBP, but this invasive investigation may be beyond the reach of most manual therapy research.[42]

A further limitation of our study was that we mainly assessed for flexion and extension pain provocation patterns. Combined movement evaluation typically seeks to identify other patterns of pain provocation (ie, lateral flexion left and right combined with flexion or extension).

Table 5. *Prevalence of positive responses during pain provocation with PAIVM tests*

	Healthy asymptomatic participants						Participants with LBP					
	PA	PA zyg. j. R	PA zyg. j. L	Lat. side P. R	Lat. side P. L	Means	PA	PA zyg. j. R	PA zyg. j. L	Lat. side P. R	Lat. side P. L	Means
L1 neutral	5.8%	5.8%	5.8%	0.0%	2.9%	4.1%	33.3%	16.6%	33.3%	23.3%	13.3%	24.0%
L1 flexion	2.9%	0.0%	0.0%	2.9%	0.0%	1.2%	20.0%	6.6%	20.0%	30.0%	16.6%	18.6%
L1 extension	11.7%	2.9%	5.8%	5.8%	0.0%	5.2%	46.6%	40.0%	33.3%	26.6%	40.0%	37.3%
L1 means	6.8%	2.9%	3.9%	2.9%	1.0%	3.5%	33.3%	21.1%	28.9%	26.6%	23.3%	26.6%
L2 neutral	0.0%	14.7%	11.7%	2.9%	2.9%	6.4%	40.0%	33.3%	40.0%	30.0%	33.3%	35.3%
L2 flexion	5.8%	8.8%	8.8%	0.0%	2.9%	5.3%	30.0%	23.3%	36.6%	30.0%	30.0%	30.0%
L2 extension	2.9%	17.6%	17.6%	5.8%	2.9%	9.4%	53.3%	36.6%	40.0%	40.0%	30.0%	40.0%
L2 means	2.9%	13.7%	12.7%	2.9%	2.9%	7.0%	41.1%	31.1%	38.9%	33.3%	31.1%	35.1%
L3 neutral	5.8%	5.8%	8.8%	8.8%	8.8%	7.6%	46.6%	30.0%	43.3%	40.0%	40.0%	40.0%
L3 flexion	8.8%	11.7%	5.8%	11.7%	11.7%	9.9%	33.3%	36.6%	50.0%	36.6%	43.3%	40.0%
L3 extension	8.8%	17.6%	11.7%	14.7%	14.7%	13.5%	60.0%	46.6%	46.6%	46.6%	43.3%	48.6%
L3 means [a]	7.8%	11.7%	8.8%	11.7%	11.7%	10.3%	46.6%	37.7%	46.6%	41.1%	42.2%	42.9%
L4 neutral	8.8%	11.7%	5.8%	20.5%	14.7%	12.3%	36.6%	40.0%	40.0%	53.3%	46.6%	43.3%
L4 flexion	14.7%	8.8%	11.7%	17.6%	23.5%	15.3%	40.0%	33.3%	43.3%	43.3%	43.3%	40.6%
L4 extension	14.7%	17.6%	17.6%	11.7%	20.5%	16.4%	50.0%	36.6%	46.6%	43.3%	50.0%	45.3%
L4 means [a]	12.7%	12.7%	11.7%	16.6%	19.6%	14.7%	42.2%	36.6%	43.3%	46.6%	46.6%	43.1%
L5 neutral	20.5%	20.5%	20.5%	23.5%	29.4%	22.9%	46.6%	40.0%	45.0%	50.0%	43.3%	45.0%
L5 flexion	11.7%	17.6%	52.9%	64.7%	58.8%	41.1%	43.3%	33.3%	43.3%	53.3%	56.6%	46.0%
L5 extension	20.5%	23.5%	20.5%	17.6%	23.5%	21.1%	46.6%	36.6%	36.6%	53.3%	50.0%	44.6%
L5 means [a,b]	17.6%	20.5%	31.3%	35.3%	37.2%	28.4%	45.5%	36.6%	41.6%	52.2%	50.0%	45.2%

Lat. side. P. R/L, lateral (transverse) pressure on spinous process right/left; *PA*, PA pressure on spinous process; *PA zyg. j. R/L*, PA pressure on zygoapophyseal joint right/left.

  [a] $P < .001$ between L5 and L1-2, L4 and L1, and L3 and L1 in the LBP group (statistical power = 1).

  [b] $P < .001$ between L5 and L1-4 in the healthy group (statistical power = 1).

Future studies could investigate the reliability of these patterns. Moreover, owing to the impossibility of blinding patients to the clinical examination and the consequential potential Hawthorne effect, patients' verbal response to pain provocative tests may have been influenced.

In manual therapy, there are various concepts for the management of LBP including MK,[16] Maitland,[30,34] Mulligan,[47] and various forms of spinal manipulation among others with different mechanisms of action.[26,30,48–52] Even if the principles of treatment vary from one method to another, the underlying principles of manual therapy are to reduce pain. It is essential that the treatment, regardless of the concept, is performed on the basis of a reliable and valid clinical examination protocol aimed to correctly classify LBP. The results of the current study using pain provocative tests provide confidence that aspects of CCR examination used in this study are valid and reliable and can therefore be used in clinical practice to direct patient management. Nevertheless, future studies are needed to confirm the value of this examination protocol. Future studies should integrate more patients with different LBP disorders (eg, in acute and chronic phases) and with more examiners.

CONCLUSION

This study demonstrated that the use of the CCR (3 positive criteria arising from active and passive pain provocative tests) was found to have good interexaminer

reliability and validity to identify the most provocative lumbar spine movement direction as well as the lumbar segmental levels of involvement.

Practical Applications
- The CCR that was developed demonstrated good reliability and validity and can be used in clinical examination of patients with LBP.
- This CCR may help clinicians in directing manual therapy treatment.

CONTRIBUTORSHIP INFORMATION

Concept development (provided idea for the research): BH, TH.

Design (planned the methods to generate the results): BH, TH.

Supervision (provided oversight, responsible for organization and implementation, writing of the manuscript): HN, CD.

Data collection/processing (responsible for experiments, patient management, organization, or reporting data): BH, TH.

Analysis/interpretation (responsible for statistical analysis, evaluation, and presentation of the results): BH, TH.

Literature search (performed the literature search): BH, TH.

Writing (responsible for writing a substantive part of the manuscript): BH, TH, CD.

Critical review (revised manuscript for intellectual content, this does not relate to spelling and grammar checking): HN, TH, CD.

REFERENCES

1. Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. Eur Spine J 2006;15(Suppl 2):S192-300.
2. Delitto A, George SZ, Van Dillen LR, et al. Low back pain. J Orthop Sports Phys Ther 2012;42:A1-A57.
3. Balague F, Mannion AF, Pellise F, Cedraschi C. Non-specific low back pain. Lancet 2012;379:482-91.
4. May S, Littlewood C, Bishop A. Reliability of procedures used in the physical examination of non-specific low back pain: a systematic review. Aust J Physiother 2006;52:91-102.
5. Hidalgo B, Gilliaux M, Poncin W, Detrembleur C. Reliability and validity of a kinematic spine model during active trunk movement in healthy subjects and patients with chronic non-specific low back pain. J Rehabil Med 2012;44:756-63.
6. 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; 75:470-85 [discussion 485–479].
7. van Kleef M, Vanelderen P, Cohen SP, Lataster A, Van Zundert J, Mekhail N. 12. Pain originating from the lumbar facet joints. Pain Pract 2010;10:459-69.
8. Kallewaard JW, Terheggen MA, Groen GJ, et al. 15. Discogenic low back pain. Pain Pract 2010;10:560-79.
9. Zusman M. Hands on, hands off? The swings in musculo-skeletal physiotherapy practice. Man Ther 2013;18:e13.
10. Vanelderen P, Szadek K, Cohen SP, et al. 13. Sacroiliac joint pain. Pain Pract 2010;10:470-8.
11. O'Sullivan P. It's time for change with the management of non-specific chronic low back pain. Br J Sports Med 2012;46: 224-7.
12. Van Dillen LR, Sahrmann SA, Norton BJ, et al. Reliability of physical examination items used for classification of patients with low back pain. Phys Ther 1998;78:979-88.
13. Maher C, Adams R. Reliability of pain and stiffness assessments in clinical manual lumbar spine examination. Phys Ther 1994;74:801-9 [discussion 809–811].
14. Harris-Hayes M, Van Dillen LR. The inter-tester reliability of physical therapists classifying low back pain problems based on the movement system impairment classification system. PM R 2009;1:117-26.

15. Riddle DL. Classification and low back pain: a review of the literature and critical analysis of selected systems. Phys Ther 1998;78:708-37.
16. McKenzie R, May S. The lumbar spine—mechanical diagnosis and therapy. 2nd. Spinal Publications, New Zealand Ltd; 2003. p. 732.
17. Kilpikoski S, Airaksinen O, Kankaanpaa M, Leminen P, Videman T, Alen M. Interexaminer reliability of low back pain assessment using the McKenzie method. Spine (Phila Pa 1976) 2002;27:E207-14.
18. Clare HA, Adams R, Maher CG. A systematic review of efficacy of McKenzie therapy for spinal pain. Aust J Physiother 2004;50:209-16.
19. 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. Arch Phys Med Rehabil 2000;81:57-61.
20. 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 (Phila Pa 1976) 2000;25:106-14.
21. 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. J Orthop Sports Phys Ther 2003;33:126-42.
22. Dankaerts W, O'Sullivan P. The validity of O'Sullivan's classification system (CS) for a sub-group of NS-CLBP with motor control impairment (MCI): overview of a series of studies and review of the literature. Man Ther 2011;16:9-14.
23. Dankaerts W, O'Sullivan PB, Straker LM, Burnett AF, Skouen JS. The inter-examiner reliability of a classification method for non-specific chronic low back pain patients with motor control impairment. Man Ther 2006;11:28-39.
24. O'Sullivan P. Diagnosis and classification of chronic low back pain disorders: maladaptive movement and motor control impairments as underlying mechanism. Man Ther 2005;10:242-55.
25. El Avd. Orthopaedic manual therapy diagnosis: spine and temporomandibular joints. 1st English ed. Sudbury, Mass.; London: Jones and Bartlett; 2010.
26. Hidalgo B, Detrembleur C, Hall T, Mahaudens P, Nielens H. The efficacy of manual therapy and exercise for different stages of non-specific low back pain: an update of systematic reviews. J Man Manip Ther 2013:1-16 in press.
27. Schneider M, Erhard R, Brach J, Tellin W, Imbarlina F, Delitto A. Spinal palpation for lumbar segmental mobility and pain provocation: an interexaminer reliability study. J Manipulative Physiol Ther 2008;31:465-73.
28. Dankaerts W, O'Sullivan P, Burnett A, Straker L, Davey P, Gupta R. Discriminating healthy controls and two clinical subgroups of nonspecific chronic low back pain patients using trunk muscle activation and lumbosacral kinematics of postures and movements: a statistical classification model. Spine (Phila Pa 1976) 2009;34:1610-8.
29. Edwards BC. Manual of combined movements: their use in the examination and treatment of musculoskeletal vertebral column disorders. 2nd ed. Oxford: Butterworth-Heinemann; 1999.
30. Maitland J. Spinal manipulation made simple: a manual of soft tissue techniques. Berkeley, Calif: North Atlantic Books; 2000.
31. McCarthy C. Combined movement theory: rational mobilization and manipulation of the vertebral column. Edinburgh: Churchill Livingstone; 2010.
32. Rubinstein SM, van Tulder M. A best-evidence review of diagnostic procedures for neck and low-back pain. Best Pract Res Clin Rheumatol 2008;22:471-82.

33. Petty NJ. Neuromusculoskeletal examination and assessment: a handbook for therapists. 3rd ed. Edinburgh: Elsevier Churchill Livingstone; 2006.

34. Hengeveld E, Banks K, Maitland GD. In: Hengeveld Elly, Banks Kevin, editors. Maitland's vertebral manipulation. 7th ed. Oxford: Elsevier Butterworth-Heinemann; 2005.

35. Haswell K, Williams M, Hing M. Interexaminer reliability of symptom-provoking active side-bend, rotation and combined movement assessments of patients with low back pain. J Manual Manipulative Ther 2004;12:11-20.

36. Horneij E, Hemborg B, Johnsson B, Ekdahl C. Clinical tests on impairment level related to low back pain: a study of test reliability. J Rehabil Med 2002;34:176-82.

37. Phillips DR, Twomey LT. A comparison of manual diagnosis with a diagnosis established by a uni-level lumbar spinal block procedure. Man Ther 1996;1:82-7.

38. Abbott JH, Flynn TW, Fritz JM, Hing WA, Reid D, Whitman JM. Manual physical assessment of spinal segmental motion: intent and validity. Man Ther 2009;14:36-44.

39. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther 2005;85:257-68.

40. Chen G, Faris P, Hemmelgarn B, Walker RL, Quan H. Measuring agreement of administrative data with chart data using prevalence unadjusted and adjusted kappa. BMC Med Res Methodol 2009;9:5.

41. Blum A, Feldmann L, Bresler F, Jouanny P, Briancon S, Regent D. Value of calculation of the kappa coefficient in the evaluation of an imaging method. J Radiol 1995;76:441-3.

42. Bosscher HA, Heavner JE. Diagnosis of the vertebral level from which low back or leg pain originates. A comparison of clinical evaluation, MRI and epiduroscopy. Pain Pract 2012; 12:506-12.

43. Kalichman L, Li L, Kim DH, et al. Facet joint osteoarthritis and low back pain in the community-based population. Spine (Phila Pa 1976) 2008;33:2560-5.

44. Maus TP, Aprill CN. Lumbar diskogenic pain, provocation diskography, and imaging correlates. Radiol Clin North Am 2012;50:681-704.

45. Hicks GE, Fritz JM, Delitto A, Mishock J. Interrater reliability of clinical examination measures for identification of lumbar segmental instability. Arch Phys Med Rehabil 2003; 84:1858-64.

46. McCullough BJ, Johnson GR, Martin BI, Jarvik JG. Lumbar MR imaging and reporting epidemiology: do epidemiologic data in reports affect clinical management? Radiology 2012; 262:941-6.

47. Exelby L. The Mulligan concept: its application in the management of spinal conditions. Man Ther 2002;7:64-70.

48. Maigne JY, Vautravers P. Mechanism of action of spinal manipulative therapy. Joint Bone Spine 2003;70:336-41.

49. Licciardone JC, Brimhall AK, King LN. Osteopathic manipulative treatment for low back pain: a systematic review and meta-analysis of randomized controlled trials. BMC Musculoskelet Disord 2005;6:43.

50. Soulvis T, Vincenzino B, Wright A. Neurophysiological effects of spinal manual therapy. In: Boyling Jeffrey D, Jull Gwendolen A, editors. Modern manual therapy, the vertebral column. In: Twomey Lance T, editor. Foreword, 3rd ed. Edinburgh: Churchill Livingstone; 2004. p. 367-79.

51. Bronfort G. Spinal manipulation: current state of research and its indications. Neurol Clin 1999;17:91-111.

52. Harvey E, Burton AK, Moffett JK, Breen A, team UBt. Spinal manipulation for low-back pain: a treatment package agreed to by the UK chiropractic, osteopathy and physiotherapy professional associations. Man Ther 2003;8:46-51.