

Original article

Construct validity of lumbar extension measures in McKenzie's derangement syndrome[☆]

Helen A. Clare^{a,*}, Roger Adams^b, Christopher G. Maher^b

^a16 Ayres Road St Ives, NSW 2075, Sydney, Australia

^bSchool of Physiotherapy, The University of Sydney, Sydney, Australia

Received 21 March 2005; received in revised form 15 May 2006; accepted 10 July 2006

Abstract

The McKenzie treatment model advocates extension-based treatments for sub-groups of low back pain (LBP) patients and an improvement in extension range is seen as a positive outcome. The treatment model states that patients who fit the McKenzie derangement classification respond faster than other patients. The validity of this treatment model and of the clinical measures of extension has not yet been established.

Fifty patients with LBP were classified as derangement ($n = 40$) or non-derangement ($n = 10$) based on a McKenzie assessment and then treated with extension procedures. Lumbar extension was measured in two positions, standing and prone, with three methods, inclinometer, Schober and finger tip to floor, on Day 1 and Day 5 of treatment. Patients completed a global perceived effect (GPE) scale on Day 5. Construct validity was tested, by comparing extension improvement and the GPE scores between the two groups. Responsiveness of the six extension measures was calculated.

All patients gained extension range however the derangement group had significantly higher GPE scores and greater improvement in extension range. The modified Schober method in standing was the most responsive method for measuring lumbar extension.

The results of this study support the measurement of lumbar extension, for patients, treated with extension procedures and provides evidence for the construct validity of one aspect of the McKenzie treatment model. The modified Schober method is the preferred protocol for a clinical setting.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Low back pain; Lumbar extension measures; McKenzie derangement classification

1. Introduction

A number of tests are used in the determination of functional capacity of patients with low back pain (LBP). While physical measures are being used less commonly to determine functional outcomes, the assessment of lumbar spinal motion is one of the tests that remains in common use (Waddell, 1991; Hazzard, 1994). There is evidence to suggest that measures of

physical impairment do not correlate highly with measures of disability in patients with LBP (Michel et al., 1977; Cox et al., 2000; Parks et al., 2003). Despite this, changes in lumbar mobility are still used as a tool to guide clinicians as to the appropriateness of the clinical intervention and as a means of monitoring progress (McKenzie, 1981; Maitland, 1986; Hahne et al., 2004).

Limitation of lumbar extension in patients with LBP is reported by McKenzie (1979), as a common finding in patients with LBP and he advocates the use of treatment procedures that move the lumbar spine into extension to assist with the resolution of symptoms (McKenzie, 1981). McKenzie and May (2003) identify subgroups of

[☆]This study was approved by the Human Research Ethics Committee of the University of Sydney.

*Corresponding author. Tel./fax: +61 02 94491027.

E-mail address: clare.ha@bigpond.com (H.A. Clare).

patients within the non-specific LBP population and utilize a classification system of derangement, dysfunction, postural and other to group patients whose symptoms respond similarly to a mechanical assessment. A high proportion of patients who fit the derangement classification demonstrate a limitation of extension range, which improves when treatment procedures that cause a reduction, abolition, or centralization of symptoms, are applied (McKenzie and May, 2003). In a recent study by Long et al. (2004) this preference for extension has been demonstrated to be present in 83% of subjects who were classified as derangement.

Patients fitting the other McKenzie subgroups may also present with a loss of extension but the recovery of both their symptoms and their range of movement is said to be slower (McKenzie, 1981).

Clinically an improvement in extension range is often seen as the patient's pain and level of function improves however measurement of change in extension range has not been well documented. McKenzie (1972) noted in a group of patients presenting with a sciatic scoliosis that lumbar extension was grossly restricted but following correction of the scoliotic deformity and centralization of the patient's symptoms their extension range was able to be restored by the patient performing extension movements in standing. Kopp et al. (1986) demonstrated that the ability of patients with herniated nucleus pulposus to achieve full passive lumbar extension was a useful predictor of a favourable response to conservative management. Smith and Mell (1987) demonstrated that prone spinal extension exercises increased lumbar extension range in a group of non-symptomatic adults, however; we were unable to locate any studies where the effect of lumbar extension exercises on patients with LBP and restricted lumbar extension had been investigated.

A narrative review of the literature (Clare, 2005) revealed four approaches that would be appropriate to measure lumbar extension in a clinical setting; inclinometer methods (20 studies), Schober variants (6 studies), fingertip to floor (FTF) method (1 study) and the use of a tape measure to determine the distance from the sternal notch to the supporting surface in the prone-press-up position (1 study). A wide variation in measurement protocols was apparent. The inclinometer protocols differed in terms of (i) whether they measured isolated lumbar or total lumbar and pelvic extension, (ii) the use of analogue or a digital inclinometer, (iii) use of a single or two inclinometers, (iv) the reference points for the placement of the inclinometer and (v) subject position in which extension was measured. The Schober variants differed in terms of the cephalad and caudad reference landmarks. Only one measurement protocol was found for both the FTF method and the sternal notch to supporting surface method.

The reliabilities reported for the various measuring devices varied considerably both in the statistic used to

describe reliability and the absolute values reported (Clare, 2005). The most optimistic values for inclinometer measures of inter-rater reliability was an intra-class correlation coefficient (ICC) of 0.93 and a coefficient variation (CV) of 2.8% whereas the least optimistic was an ICC of 0.16 and a CV of 28.5%. The reliabilities reported for the various Schober methods were more consistent ranging from an ICC of 0.65–0.97 for intra-rater and from 0.54 to 0.94 for inter-rater. A single study reported inter-rater reliability of ICC = 0.96 for the finger tip to floor method of measuring extension (Hahne et al., 2004). A single study utilizing the measurement of the distance from the sternal notch to the supporting surface when the patient performed extension in lying exercises reported inter-rater reliability of ICC = 0.85 (Bandy and Reese, 2004).

Few studies evaluated criterion validity. Correlations with radiographic assessment and inclinometer measures varied from 0.60 to 0.76 and a correlation of 0.49 with a modified Schober measure has been reported (Clare, 2005).

Longitudinal construct validity (internal and external responsiveness) of extension range has not been widely studied. Pengel et al. (2004) noted that extension range of motion, measured with a single inclinometer, was the most responsive of the range of motion measures evaluated. Interestingly the responsiveness was similar to that of the Roland Morris disability scale but less than that of the Patient Specific Functional Scale or pain numerical rating scale (Pengel et al., 2004).

The primary aims of the study therefore were to investigate whether:

- Lumbar extension range improves with the use of McKenzie's extension treatment procedures.
- In patients treated with the use of McKenzie's extension treatment procedures, those classified as derangement show a greater improvement in extension range and report greater recovery, than those in all other McKenzie classifications.

The secondary aims of the study therefore were

- to compare the responsiveness of the six extension measures and
- to investigate the relationship between six simple clinical measures of extension range of motion.

2. Methods

2.1. Subjects

Consecutive new patients attending a private physiotherapy clinic for LBP were invited to participate in

the study with the data being collected over a 6-month period. The criteria for inclusion were that they were currently experiencing non-specific mechanical LBP of any duration, with or without radiation to the leg. Patients were excluded if they had non-mechanical spinal pain, if medical conditions limited their ability to perform the McKenzie procedures or if they had a history of spinal surgery. All subjects gave written consent prior to participating.

Information was collected from the subjects regarding their gender, age, location of symptoms, duration of symptoms, previous history of LBP, pain intensity and functional status.

2.2. Procedure

The patients underwent a standardized McKenzie assessment, and were classified as either, derangement or non-derangement, according to the operational definitions described by McKenzie and May (2003) (Appendix A). All the assessments were performed by the one examiner, (researcher HC), who had 18 years experience using the McKenzie method. All the patients classified as derangement demonstrated a directional preference for extension during the assessment and based on the assessment findings extension procedures were the treatment of choice for the non-derangement patients.

At the conclusion of the assessment the measures of extension range were taken.

Standard protocols for the measurement procedures were defined and practiced by the therapist who performed all the measurements (Appendix B). With the subject in relaxed standing and the therapist kneeling behind the patient the left and the right posterior superior iliac spines (Dimples of Venus) were identified and a line was drawn between them. The point where this line was bisected by the spine was marked and a second mark was made on the spine at a distance of 15 cm above this.

The subject was then taught how to perform extension in standing with their feet placed shoulder width apart, with a minimum of three extension movements being performed as practice. Three measures of extension in standing were then taken (finger tip to floor, single inclinometer, and modified Schober) with the measures taken in a random order. The subject was then taught how to perform extension in lying as described by McKenzie (1981) again with a minimum of three being performed to familiarize the subject with the movement. Two measures of extension were taken of extension in lying (inclinometer and modified Schober).

The subjects were then provided with the extension treatment procedures deemed appropriate by the treating physiotherapist for their classification and clinical presentation. The treatment procedures included the

following: extension in lying exercises, extension in lying with over-pressure, extension mobilization, extension in standing exercises, and extension in sitting with maintenance of the lumbar lordosis by the use of a lumbar support (McKenzie and May, 2003). All subjects were provided with a home program of extension exercises and advice. Modalities, drug therapy, corsets, spinal manipulative therapy other than extension mobilization were not provided.

After five treatment sessions the extension measures were repeated in the same manner and the subjects completed a 7-point global perceived effect (GPE) scale which ranged from 1—completely recovered to 7—vastly worse (Beurskens et al., 1996).

This study was approved by the Human Research Ethics Committee of the University of Sydney.

2.3. Data analysis

To evaluate the hypothesis that lumbar extension range improves with the use of McKenzie's extension treatment procedures, the Day 1 and Day 5 extension measures were compared using paired-samples *t*-tests.

To evaluate the hypothesis that in patients treated with the use of McKenzie's extension treatment procedures, those classified as derangement would show a greater improvement in extension range than those classified as non-derangement, linear regression analyses were conducted to predict the Day 5 extension score based upon the Day 1 extension score and type of McKenzie syndrome. The hypothesis would be supported if the McKenzie syndrome was a significant predictor of Day 5 extension range.

To evaluate the hypothesis that with patients treated with the use of McKenzie's extension treatment procedures, those classified as derangement would report greater recovery (GPE score) than those classified as non-derangements, GPE scores for the two groups were compared using a Mann-Whitney *U*-test.

To evaluate the relationship between six simple clinical measures of extension range of motion, Pearson's product moment correlations (PMC) were calculated. To describe the responsiveness of the six extension measures the effect size (ES) and standardized response mean (SRM) were calculated. The ES was the mean change divided by the baseline standard deviation and the SRM was the mean change divided by the standard deviation of the change amounts. We calculated 84% confidence intervals for direct comparison of ES. These were chosen because non-overlapping 84% confidence intervals are equivalent to a *Z* test of means at the 0.05 level (Tryon, 2001). Lastly the GPE scores and the change scores for each of the extension measures were correlated using Pearson's Product Moment Correlation with 95% CI.

Table 1
Baseline characteristics of the subjects

Characteristics of subjects	All subjects (<i>n</i> = 50)	Derangements (<i>n</i> = 40)	Non-derangement (<i>n</i> = 10)
Age (years)	50.3 (14.54)	48.32 (13.62)	58.36 (16.06)
Female, gender	50%	50%	50%
Pain intensity (VAS cm)	5.10 (1.64)	4.90 (1.62)	5.75 (1.64)
Quebec disability score	54.8 (14.32)	53.20 (13.17)	61.32 (17.56)
Previous episodes LBP % yes	92%	93%	90%
% derangement	80%	N/A	N/A
Location of symptoms			
Central back pain	38%	40%	30%
Radiation to the knee	40%	40%	40%
Radiation below the knee	22%	20%	30%
Duration of symptoms ^a			
Acute (<7days)	14%	17.5%	0%
Subacute (>7 days–7 weeks)	24%	27.5%	10%
Chronic (>7 weeks)	62%	55%	90%

Data for continuous variables are means (standard deviations) and categorical variables are percentages.

^aSpitzer et al. (1987).

Table 2
Change in extension range from Day 1 to Day 5

Extension measure	Day 1 (mean (SD))	Day 5 (mean (SD))	Extension improvement (mean (SD))		
			Total group (<i>n</i> = 50)	Derangement (<i>n</i> = 40)	Non- derangement (<i>n</i> = 10)
Inclinometer standing (deg)	18.2 (5.32)	26.5 (6.16)	8.3 (3.59)	9.13 (3.56)	5.00 (0.00)
Inclinometer prone (deg)	23.9 (6.00)	32.5 (8.18)	8.56 (5.29)	9.82 (5.06)	3.50 (2.42)
Schober standing (cm)	1.24 (0.49)	2.12 (0.623)	0.88 (0.37)	0.97 (0.36)	0.50 (0.00)
Schober prone (cm)	1.80 (0.62)	2.64 (0.79)	0.84 (0.46)	0.98 (0.39)	0.30 (0.26)
Finger tip to floor standing (cm)	64.2 (5.32)	60.1 (5.54)	4.14 (1.97)	4.52 (2.17)	2.00 (0.66)
Finger tip to floor standing (cm) (normalized for height)	0.38 (0.02)	0.36 (0.03)	0.025 (0.01)	0.03 (0.01)	0.01 (0.01)

In the group of 50 patients extension range improved with treatment (all $P < 0.001$). Regression analyses revealed that the improvement was greater for the derangement group (all $P < 0.001$).

3. Results

The baseline characteristics of the subjects who participated in the study are described in Table 1.

The group mean and standard deviation for Days 1 and 5 extension scores are shown in Table 2. All six measures of extension improved significantly from Days 1 to 5 (all $P < 0.001$).

The improvements in extension range for the derangement and non-derangement groups are shown in Table 2. For each regression analysis the baseline score and McKenzie group (derangement versus non-derangement) were significant predictors of the Day 5 score (all $P < 0.001$). These results show that in patients treated with the use of McKenzie's extension treatment procedures, those classified as derangement achieve a

significantly greater improvement in extension range than those in all other McKenzie classifications.

In patients treated with the use of McKenzie's extension treatment procedures, those classified as derangement also reported greater perceived recovery than other subjects. The median GPE score at Day 5 in the derangement group was 2, i.e. much improved, (range 1–2) and in the non-derangement groups it was 3, i.e. slightly improved (range 3–3). Results from a Mann–Whitney *U*-test revealed that the derangement group had greater GPE scores than the non-derangement group (all $P < 0.001$).

The PMC results revealed a complex relationship between the six extension measures (Table 3). The baseline inclinometer and modified Schober measures performed in standing and prone were highly correlated

Table 3
Pearson product moment correlations between the 6 extension measures

	Inc/std	Inc/prone	Mod Sch std	Mod Sch prone	FTF	FTF norm
Inc/std	1					
Inc/prone	0.82 (0.65)	1				
Mod Sch std	0.70 (0.57)	0.60 (0.61)	1			
Mod Sch prone	0.74 (0.58)	0.79 (0.66)	0.84 (0.82)	1		
FTF	0.02 (-0.48)	0.10 (-0.48)	-0.12 (-0.50)	-0.04 (-0.50)	1	
FTFnorm	-0.04 (-0.38)	-0.02 (-0.40)	-0.13 (-0.40)	-0.10 (-0.52)	0.84 (0.83)	1

The baseline correlations are non-bold, the correlations between the change scores are bold and are in brackets.

Inc, inclinometer, Mod Sch, modified schober, std, standing, FTF, finger tip to floor, FTFnorm, finger tip to floor normalized.

Table 4
Responsiveness of the 6 extension measures

	Effect size (84% CI)	Standardised response mean (84% CI)	Correlation (95% CI)
Inclinometer standing	1.56 (1.42–1.70)	2.31 (2.12–2.51)	0.58 (0.36–0.74)
Inclinometer prone	1.43 (1.25–1.60)	1.62 (1.42–1.82)	0.51 (0.27–0.69)
Modified Schober standing	1.81 (1.66–1.96)	2.37 (2.17–2.56)	0.60 (0.39–0.75)
Modified Schober prone	1.35 (1.20–1.49)	1.84 (1.64–2.04)	0.66 (0.47–0.80)
Finger tip to floor standing	0.78 (0.70–0.86)	2.10 (1.88–2.32)	0.54 (0.30–0.71)
Finger tip to floor standing (normalized for height)	1.07 (0.96–1.18)	1.94 (1.74–2.14)	0.51 (0.26–0.69)

($r = 0.60$ – 0.84), however; these measures were unrelated to the baseline FTF scores (absolute value of all correlations less than 0.15). The correlations between the change scores for the inclinometer and the modified Schober measures were slightly lower ($r = 0.57$ – 0.82) than observed for the baseline correlations. Contrary to the results for baseline scores, changes in the FTF measures were moderately correlated with changes in other scores ($r = -0.38$ to -0.52).

Three measures of responsiveness were calculated for the six measures and are provided in Table 4. The modified Schober measurement taken in standing had both the largest ES and the highest SRM. In addition, its correlation was second to the modified Schober measurement taken in the prone position.

4. Discussion

The results of this study support the aspect of the McKenzie treatment model that was evaluated. Consistent with the treatment model, patients treated with an extension approach improved both their extension range and amount of reported improvement in their condition. Moreover, those classified as derangement showed greater improvement, as seen in measurements of extension range and GPE ratings. This is in keeping with the findings of Werneke et al. (1999) who demonstrated that patients whose symptoms centralize

(derangement) improve faster and have better outcomes than those patients whose symptoms did not centralize (non-derangement).

The concept of sub-classification of patients within the spectrum of non-specific LBP is supported by this study. Using well-defined classification criteria (McKenzie and May, 2003) the patients were classified into derangements or non-derangements. Although both groups of patients were treated with extension procedures the patients with derangements recovered extension faster than the patients with non-derangement and reported better satisfaction with treatment. Identification of the derangement sub-group appears to allow prediction of their short-term outcomes, both by a physical measure and with a patient orientated measure. Waddell et al. (1992) found only a weak association between pain, disability and physical impairment in a group of patients with chronic LBP and others have advocated that more emphasis should be placed on change in pain and disability scores than on change in physical impairment when gauging treatment outcomes (Deyo et al., 1994; Pengel et al., 2004). However in specific sub-groups of patients with LBP who present with a loss of extension the physical measure of lumbar extension appears to be a useful tool in determining short-term outcomes along with measures of function. Further research needs to determine if these improvements in physical measures are correlated with long-term functional improvements.

The degree of extension loss forms part of the physical assessment of the LBP patient as described by McKenzie (1981), however observation rather than measurement of extension is advocated. The results of this study suggest that lumbar extension can be measured effectively in a clinical setting with an inexpensive measuring tool and that the measurement can provide valuable clinical information. Accordingly, we would advocate that practitioners who use the McKenzie approach make an objective measurement of lumbar extension rather than a subjective one through observation.

Review of the literature revealed a number of different ways of measuring lumbar extension, and the results of this study show that the different methods are not interchangeable. The inclinometer measures and the modified Schober measures appeared to be measuring similar things (both are measuring isolated lumbar extension as they are ‘on the skin’ measures) however the finger tip to floor method was very distinct with minimal relationship to the other methods, especially at baseline.

At present, there is not yet agreement on the optimal design and analysis to evaluate responsiveness (Stratford et al., 1996). Accordingly, three of the most commonly used approaches were employed. Considering all three approaches, the most responsive measure was the modified Schober in standing followed by the inclinometer in standing. Whilst there was little difference between these two, the FTF method was considerably less responsive. For example the ES for the modified Schober in standing was 1.81 (1.66–1.96) and for FTF 0.78 (0.70–0.86). This result argues against the use of the FTF method.

Over the past decade there has been a move away from the previous use of range of movement as an outcome measure and there is now more emphasis on self-report measures of pain and disability for the determination of outcomes following treatment (Deyo et al., 1994). The results of this study suggest that in specific sub-groups of patients who present with a loss of lumbar extension and are treated with extension procedures it would be appropriate to supplement the pain and disability measures with the measurement of extension. An advantage that extension measures have over disability measures such as the Quebec LBP disability scale (Kopeck et al., 1995) is that they can be re-administered over much shorter time intervals. It would be appropriate to measure extension pre and post a single treatment session but not disability using the Quebec scale, as the patient would not have had the opportunity to perform the functional activities described in the Quebec scale items.

The limitations of this research project include the lack of blinding of the therapist taking the extension measures to the classification of the patient, that there was a single treating therapist and a modest number of

patients. Now that a responsive measure of lumbar extension has been determined the study needs to be replicated on a larger LBP population, and in a variety of health care settings. Multiple, McKenzie-trained therapists should be utilized to perform the assessments and treatments with an independent examiner performing the extension measures. A further limitation of the study was that there was no measurement of pain and disability on Day 5. This data may have assisted in determining outcomes. In future studies the extension measures and pain and disability measures should be repeated both at short- and long-term follow-ups.

5. Conclusion

The results of this study provide evidence for the construct validity of one aspect of the McKenzie treatment model. Extension range improved in all the patients who were treated with extension procedures, but a greater improvement in extension range and in GPE was demonstrated in the sub-group of patients classified as derangement. These findings provide support for the measurement of lumbar extension along with pain and disability for patients who are treated with extension procedures. Of six extension measures, the modified Schober method was supported as the preferred method for clinical settings.

Acknowledgements

Christopher Maher’s research fellowship is funded by Australia’s National Health and Medical Research Council.

Appendix A. Operational definition (McKenzie and May, 2003)

A.1. Derangement

Centralization: In response to therapeutic loading strategies pain is progressively abolished in a distal to proximal direction, AND each progressive abolition is retained over time, until all symptoms are abolished, AND if back pain only is present this moves from a widespread to a more central location and then is abolished OR pain is decreased and then abolished during the application of therapeutic loading strategies.

The change in pain location, or decrease or abolition of pain remain better, and should be accompanied or preceded by improvements in the mechanical presentation (range of movement and/or deformity).

NB: In this study only derangement cases where there was a directional preference for extension were included.

Appendix B. Protocol for measurement procedures

B.1. Finger tip to floor (FTF)

The subject is asked to stand with the feet placed shoulder width apart. They are then asked to arch backwards as far as they can, reaching their fingertips down the backs of their legs whilst keeping their knees straight. The rater checks that the finger tips of each hand are at equal levels, marks the fingertip level with a skin marker. The perpendicular distance between the marked point and the floor is then measured with a wooden ruler.

B.2. Modified Schober: standing/lying

With the subject in relaxed standing the rater marks the level of the bisector of the line running between the left and the right posterior superior iliac spines and then makes a mark on the spine 15 cm directly above the first. The patient is asked to bend backwards as far as possible whilst keeping their knees straight/by pushing up as high as they can.¹ Whilst in this position the distance between the two marks is measured again with the tape resting on the skin.

B.3. Single inclinometer: standing/lying

A single straight-edged inclinometer is placed on the spinous process of T12 and a base reading is taken. The subject is asked to arch backwards as far as possible whilst keeping their knees straight/by pushing up as high as they can.¹ Whilst in this position a second reading is taken from the inclinometer. The range of extension is the difference between the two measures.

References

- Bandy WD, Reese NB. Strapped versus unstrapped technique of the prone press-up for measurement of lumbar extension using a tape measure: differences in magnitude and reliability of measurements. *Archives of Physical Medicine and Rehabilitation* 2004;85:99–103.
- Beurskens A, de Vet HC, Kokee AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain* 1996;65:71–6.
- Clare HA. Evaluation of the McKenzie method. PhD thesis, USyd. 2005. p. 117–9 [Chapter 8].
- Cox ME, Asselin S, Gracovetsky SA, Richards MP, Newman NM, Karakusevic V, et al. Relationship between functional measures and self-assessment in non-acute LBP. *Spine* 2000;25:1817–26.
- Deyo R, Andersson G, Bombardier C, Cherkin D, Keller R, Lee C, et al. Outcome measures for studying patients with low back pain. *Spine* 1994;19:2032S–6S.
- Kopec J, Esdaile J, Abrahamowicz M, Abenhaim L, Wood-Dauphinee S, Lamping D, et al. The Quebec back pain disability scale. Measurement properties. *Spine* 1995;20:341–52.
- Kopp J, Alexander H, Turocy R, Levrini M, Lichtman D. The use of lumbar extension in the evaluation and treatment of patients with acute herniated nucleus pulposus. *Clinical Orthopaedics* 1986;202:211–8.
- Hahne A, Keating J, Wilson S. Do within-session changes in pain intensity and range of motion predict between-session changes in patients with low back pain? *Australian Journal of Physiotherapy* 2004;50:17–23.
- Hazzard R, Williams M, McKenzie R. Reliability of three methods for measuring prone lumbar extension. *New Zealand Journal of Physiotherapy* 1994;1:11–3.
- Long A, Donelson R, Fung T. Does it matter which exercise? *Spine* 2004;29:2593–602.
- Maitland G. *Vertebral manipulation*, fifth ed. London: Butterworths; 1986. p. 144 [Chapter 9].
- McKenzie RA. Manual correction of sciatic scoliosis. *New Zealand Medical Journal* 1972;76:194–9.
- McKenzie RA. Prophylaxis in recurrent low back pain. *New Zealand Medical Journal* 1979;89:22–3.
- McKenzie RA. *The lumbar spine: mechanical diagnosis and therapy*, first ed. Waikanae: Spinal Publications; 1981. p. 39 [Chapter 7].
- McKenzie RA, May S. *The lumbar spine: mechanical diagnosis and therapy*, second ed. Waikanae: Spinal Publications; 2003. p. 557 [Chapter 23].
- Michel A, Kohlmann T, Raspe H. The association between clinical findings on physical examination and self-reported severity in back pain: results of a population based study. *Spine* 1977;22:296–303.
- Parks KA, Crichton KS, Golford RJ, McGill SM. A comparison of lumbar range of motion and functional ability scores in patients with low back pain. *Spine* 2003;28:380–4.
- Pengel L, Refshauge K, Maher C. Responsiveness of pain, disability, and physical impairment outcomes on patients with low back pain. *Spine* 2004;29:879–83.
- Smith R, Mell D. Effects of prone spinal extension exercise on passive lumbar extension range of motion. *Physical Therapy* 1987;67:1517–21.
- Spitzer W, Le Blanc FE, Dupuis M, et al. Scientific approach to the activity assessment and management of activity-related spinal disorders. *Spine* 1987;12:S1–S55.
- Stratford P, Binkley JM, Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. *Physical Therapy* 1996;76:1109–23.
- Tryon W. Evaluating statistical difference, equivalence, and indeterminacy using inferential confidence intervals: an integrated alternative method of conducting null hypothesis statistical tests. *Psychological Methods* 2001;6:371–86.
- Waddell G, Allan D, Newton M. Clinical evaluation of disability in low back pain. In: Frymoyer J, editor. *The adult spine: principles and practice*. first ed. New York: Raven Press; 1991. p. 155–68 [Chapter 5].
- Waddell G, Somerville D, Henderson I, Newton M. Objective clinical evaluation of physical impairment in chronic low back pain. *Spine* 1992;17:617–27.
- Werneke M, Hart D, Cook D. A descriptive study of the centralization phenomenon: a prospective analysis. *Spine* 1999;24:676–83.

¹The pelvis was not stabilized by a belt when the extension was performed in prone.