Predictor Variables for a Positive Long-Term Functional Outcome in Patients with Acute and Chronic Neck and Back Pain Treated with a McKenzie Approach: A Secondary Analysis

STEPHEN MAY¹, ERIC GARDINER², STEVE YOUNG³, JENNIFER KLABER-MOFFETT⁴

Back and neck pain are common symptoms in the general population and in those seeking healthcare¹⁻⁵. In both conditions, symptoms are frequently protracted and recurrent episodes are common^{2-4,6}. The majority of randomized controlled trials have produced only “trivial” treatment effects⁷, and international guidelines provide only

bland and generalized recommendations that are of limited use in the clinic⁸. It has been suggested that classification into sub-groups prior to treatment would enhance treatment efficacy and that it is a fertile ground for research⁹. Indeed, there is accumulating evidence that treatment based on classification or focused on subgroups of the back pain population

results in superior outcomes¹⁰⁻¹⁶. These studies principally investigated the McKenzie¹⁷⁻²⁰ or Treatment-Based Classification systems¹⁵.

The McKenzie method is a commonly used classification-based approach for the management of spinal pain²¹⁻²³. This is a biomechanical approach based on classification into non-specific mechanical syndromes that guide the prescription of specific directional exercises, such as extension or flexion exercises¹⁷⁻²⁰. These exercises are determined during assessment by positive symptom responses such as centralization or a decrease or abolition of pain. Centralization is the abolition of distal pain in response to repeated movements or sustained postures.

Evidence for the reliability and prognostic validity of centralization has been demonstrated in a number of studies²⁴. Evidence for the importance of specific exercises linked to directional preference has also been demonstrated¹⁰. Two systematic reviews found that there was limited evidence for the effectiveness of the McKenzie method, but that overall the literature was too limited to make definitive conclusions^{25,26}. Another systematic review evaluated the effectiveness of physical therapy-directed exercise interventions after patients had been classified using symptom response methods, with

ABSTRACT: A cognitive behavioral approach was previously compared to a biomechanical approach (the McKenzie method) for the treatment of patients with back and neck pain in a randomized trial. Few differences between the treatment interventions were found. The aim of this secondary analysis was to determine if any clinical characteristics distinguished those patients who responded best to the McKenzie approach. Treatment success was defined as 50% reduction in original functional disability scores (Roland-Morris Disability Questionnaire or Northwick Park Neck Pain Questionnaire); failure to achieve this was defined as treatment failure. A *liberal* definition of success was 50% improvement retained at either 6 or 12 months, whereas a *strict* definition of success was 50% improvement at both 6 and 12 months. Ten variables were screened by univariate regression analysis to see if they predicted success. Any significant variables ($P < 0.1$) underwent multiple regression analysis. Only 21 and 16 patients out of 102 were deemed treatment successes according to the liberal and strict definitions, respectively. With the liberal definition, only centralization ($P = 0.065$), spine region (back rather than neck pain) ($P = 0.089$), and duration of pain ($P = 0.001$) emerged as predictors from the univariate regression analysis. With the strict definition, only the latter two variables emerged: spine region ($P = 0.026$) and duration of pain ($P < 0.01$). All these variables were retained in the multiple regression analysis. In this study, duration of pain was the strongest predictor of success, although back pain and centralization had some predictive ability.

KEYWORDS: Low Back Pain, McKenzie, Multivariate Analysis, Neck Pain, Secondary Analysis, Solution-Finding Approach

¹Senior Lecturer in Physiotherapy, Faculty of Health and Wellbeing, Sheffield Hallam University, Sheffield, UK.

²Statistician, Institute of Rehabilitation, University of Hull, Hull, UK.

³Consultant Physiotherapist, Physiotherapy Department, Royal Oldham Hospital, Pennine Acute Hospitals NHS Trust, UK.

⁴Professor of Rehabilitation and Therapies, Institute of Rehabilitation, University of Hull, Hull, UK.

Address all correspondence and requests for reprints to: Stephen May, s.may@shu.ac.uk

4 out of 5 studies relating to the McKenzie method¹¹. The review concluded that exercise programs implemented in line with patient responses had significantly better results than the pragmatic control comparisons.

Most randomized controlled trials do not differentiate patients with spinal pain in any way, and randomize without consideration as to the suitability of the treatment. However, there is accumulating evidence that sub-groups of patients with back pain respond better to certain interventions. For instance, patients who have a mechanically determined directional preference are treated more effectively when matched to their directional preference than if treated with alternate or non-specific exercises¹⁰. In addition, there is some evidence that certain clinical characteristics predict a better response to manipulation or to stabilizing exercises for back pain¹²⁻¹⁴. It has been argued that classification-led treatment will improve treatment outcomes, and progress beyond the trivial effects seen in many trials^{10,15,16}. If this is the case, it is important to begin to determine what clinical characteristics differentiate those who respond well to an intervention from those who do not.

A randomized controlled trial was conducted that compared the McKenzie method with a cognitive behavioral approach known as the Solution-Finding Approach (SFA) to determine which was more effective at improving coping strategies and reducing pain and disability²⁷. The aim of this secondary analysis of the data from that trial was to determine if there were any clinical characteristics that differentiated patients who responded well to the McKenzie method from patients who did not respond well.

Methods

Original Trial

The methods of the original trial are described in detail elsewhere²⁷ and are briefly summarized here. Between March 2003 and July 2004, 315 patients were recruited (219 with back pain and 96 with neck pain); patients were referred by GPs to National Health Service

(NHS) physiotherapy departments in hospitals in West and East Yorkshire, UK. Patients were randomly allocated to McKenzie treatment or the SFA, then further randomized to receive an educational booklet (*The Back Book* or *The Neck Book* as appropriate) or no booklet. Patients were followed up by postal questionnaire at 6 weeks, and then again at 6 and 12 months. Treating therapists provided both interventions. All physiotherapists had completed parts A to D training courses in the McKenzie method and were given a training session over a day and a half in the use of the SFA and provided with a manual.

The primary outcomes were the Tampa Scale for Kinesiophobia (TSK) Activity Avoidance subscale and the Roland-Morris Disability Questionnaire (RDQ) or the Northwick Park Neck Pain Questionnaire (NPQ); however, a number of other outcomes were used for measuring general health and psychological well-being. Both groups improved somewhat over time, averaging around 2.2 on the TSK Activity Avoidance subscale, and around 2.7 on the RDQ or 2.8 on the NPQ by 6 weeks, with similar modest improvements in most secondary outcome measures. Improvements at 6 and 12 months respectively averaged 2.9 and 2.6 (TSK Activity Avoidance), 3.9 and 3.9 (RDQ), and 5.0 and 5.3 (NPQ).

There were no significant differences between the groups except in two minor instances of limited clinical significance. From the original paper, it was estimated that mean improvement from baseline functional disability measures ranged from 20% to about 32%; these mean differences were likely to conceal considerable individual variability.

Secondary Analysis

This secondary analysis sought to discriminate which patients responded to the McKenzie method better than average. Treatment success was *a priori* defined as a 50% reduction in RDQ or NPQ scores from baseline to 6 weeks, and retained at 6 or/and 12 months. We used a liberal definition of success as maintained at *either* 6 or 12 months, and

a strict definition as maintained at *both* 6 and 12 months. A 50% reduction in disability measures has been previously justified in the use of the formulation of clinical prediction rules¹²⁻¹⁴. Patients with less than 50% reduction in disability measures were deemed as treatment failures. Treatment successes and treatment failures within the McKenzie group were compared regarding certain demographic and clinical factors, namely age, gender, pain duration, pain location, spine region, McKenzie classification, therapeutic force, and centralization/abolition of symptoms.

Data Analysis

Univariate logistic regression analysis was used as a screening tool to examine the relationship between each of the putative predictors and the outcome (success or failure) at 6 weeks and at 6 or/and 12 months. The number of variables considered was restricted, as predictive factor testing requires an adequate number of outcomes per individual predictor; at least 10 cases per factor has been recommended²⁸. Ignoring this rule-of-thumb is likely to lead to *overfitting* (a model will be produced that appears to fit the sample data well but performs poorly on new data²⁹).

Two of these factors were dependent on evaluation by the treating therapists; both classification³⁰⁻³² and centralization³¹⁻³⁴ have been shown to be reliably evaluated by trained therapists with kappa scores greater than 0.6, the level of reliability coefficient said to be appropriate for a clinical prediction rule³⁵. Consistent with Hicks et al¹⁴ and Jellema et al³⁶, the level of significance for the univariate screening regressions was set at $P=0.2$, assessed by likelihood ratio tests; more stringent significance levels can lead to the exclusion of potentially useful predictor variables. Predictor variables found to be significant according to this criterion were entered into a multiple logistic regression model. For this study, variables with likelihood ratio test P -values of less than 0.1 were retained in the multiple logistic regression models. The Hosmer and Lemeshow statistic was used as a simple way to check whether the logistic regression

models predicted poorly for any groups of patients, when grouped together into similar probabilities of treatment success. Predicted probabilities of success from the final logistic regression models chosen were obtained from the mathematical equations for these models.

Results

Of the 161 patients randomized to the McKenzie group in the RCT, 102 had sufficient data available for treatment

success or failure to be defined. According to the pre-defined criteria, there were 21 (20.6%) treatment successes according to the liberal definition, and 16 (15.7%) cases according to the strict definition. Additionally, there was missing data for some of the predictor variables eliminating some cases; between 91 and 102 cases were usable for each predictor (Table 1). Only pain duration, spinal region, and centralization response merited inclusion in the multiple logistic regression analyses. In the anal-

yses for both the liberal and strict definition of treatment success, all entered variables remained significant (Table 2). The Hosmer and Lemeshow test statistics were: $\chi^2 = 2.77$; d.f.=6; $P=0.837$ (liberal definition) and $\chi^2 = 2.132$; d.f.=2; $P=0.344$ (strict definition), both indicating adequate model fits. However, the final models to predict probability of success had limited clinical applicability as even the best subgroup had less than a 70% chance of success (Tables 3 and 4). For the liberal definition, the chance of success ranged from about 69% for patients with back pain, pain for less than 12 weeks, and with centralization/abolition of pain to around 3% for patients with neck pain, pain of more than 12 weeks duration, and with no centralization/abolition (Table 3).

TABLE 1. Results of univariate logistic regression analyses for demographic and clinical factors.

Factor	Categories	Results		
		N	PI*	P2*
Age	Continuous data	102	0.315	0.296
Gender	Male	102	0.988	0.947
	Female			
Pain duration	Sub-acute (< 12 weeks)	100	0.001	<0.001
	Chronic (> 12 weeks)			
Spine or referred pain	QTF 1 and 2	97	0.547	0.414
	QTF 3 and 4			
Spine region	Neck pain	102	0.089	0.026
	Back pain			
Classification	Derangement, dysfunction, or postural mechanical syndromes**	102	0.557	0.989
	Other			
Therapeutic force	Patient forces only	91	0.699	0.593
	Additional therapist force			
Pain change status	Centralization/abolition of pain	102	0.065	0.642
	No centralization/abolition of pain			

QTF = Quebec Task Force 1 = spine pain; 2 = additional pain radiating to knee or elbow; 3 = referred pain below the knee or elbow; 4 = 3 plus neurological signs and symptoms

* PI = liberal definition; P2 = strict definition (see text)

Bolded figures = significant

** = Nomenclature of McKenzie's mechanical syndromes

TABLE 2. Results from multiple logistic regression analyses (N=100).

Variable	P-value	Odds ratio	95% CI Odds ratio
Centralization*	0.082	2.740	0.882 to 8.475
Back region:			
Liberal definition	0.038	3.636	0.985 to 13.419
Strict definition	0.01	3.547	0.998 to 12.611
< 12 weeks duration:			
Liberal definition	<0.001	7.255	2.365 to 22.262
Strict definition	<0.001	6.826	2.308 to 20.185

*Liberal definition only

Discussion

In this secondary analysis of a randomized controlled trial²⁷, there were few variables that predicted treatment success, defined as 50% reduction in disability scores at 6 weeks and maintained at 6 or 12 months (liberal definition) or maintained at 6 and 12 months (strict definition). Even by the liberal definition, treatment success was uncommon. Pain duration (less than 12 weeks) and spinal region (back pain) were retained in both definitions, and centralization or abolition of pain was retained in the liberal definition as well. Variables identified as important in the univariate screening were retained in the multivariate analyses. As the overall success rate was low, it was not surprising that variables predicted only a limited chance of success in any subgroup. The most important factor to emerge as a strong predictor was duration of pain. For either definition of success, patients with pain duration less than 12 weeks had odds of success about 7 times greater than patients with report of longer pain duration. Patients with back pain had odds of success about 3.5 times greater than those with neck pain. Patients with centralization or abolition of pain had odds of success about 2.7 times greater than those without these symptom responses, according to the liberal definition of success only.

Chronic spinal pain has traditionally been deemed as more difficult to treat so this finding was not new. This implies that patients with back or neck problems should be seen within 3 months; otherwise, the chance of success is considerably reduced. The finding that back pain is treated more successfully than neck pain has not been noted before and is in contrast to work by Werneke et al³⁴, who found that improvements in all spinal patients who demonstrated centralization or partial centralization were not significantly different between neck and back pain but were significantly different from non-centralizers. Furthermore, the lack of value of classification and the limited value of centralization to predict success also is in contrast to a number of previous studies^{10,11,15,16,34}.

Limitations

There are several weaknesses to this secondary analysis. Although 163 patients were randomized to the McKenzie group, of whom 132 completed follow-up, partial missing data meant that only 102 of the original cohort were included in the analysis. The definition of treatment success used had been previously

justified¹²⁻¹⁴ but applied poorly in this particular setting with only 20% or 16% being deemed treatment successes depending on the definition used. The definition of success was previously applied in a sub-acute and acute population whereas the majority of our population was chronic. Other definitions of success might have captured a larger population and been more discriminating. We tested this speculation by re-examining the data with a definition of success of 25% improvement in functional disability. Although post-hoc, this was felt justifiable as previously the minimal level of detectable change had been defined as 4 to 5 points for moderate disability on the RDQ³⁷, which amounts to 17% to 21% improvement; and 21% improvement for a different neck functional outcome, the Neck Disability Index³⁸. This definition of success did, of course, result in more treatment successes: 41 (40%) with the liberal definition and 34 (33%) with the strict definition. However, statistical modeling to identify predictors was less illuminating: only pain duration ($P<0.001$) and centralization ($P=0.061$) were significant in the univariate analysis under the liberal definition, and only pain duration ($P<0.001$) under the strict defini-

tion. These variables were retained in the multiple regression analysis.

In our main data analysis, the best scenario was back pain of less than 12 weeks duration with centralization or abolition of pain; such a patient is predicted to have a 69% probability of success. A patient with back pain less than 12 weeks duration and no centralization is predicted to have a 45% probability of success, and one with neck pain less than 12 weeks duration with centralization or abolition of pain is predicted to have a 38% probability of success. All other category combinations had probability of success less than 25%. The models therefore had limited clinical applicability, but it is unclear if this was because unknown important variables were not included or the definition of success was inappropriate. The above probabilities of success were all with the liberal definition of success. It is important to note that with the strict definition of success, back pain of less than 12 weeks had a predicted probability of success of less than 50%, with all other models being extremely weak. The strict definition was obviously a better measure of success as the improvement in functional disability was maintained at both medium- and long-term. Perhaps a 50% improvement was an unrealistic target for our study population.

Numerous other baseline clinical and demographic data could have been included in the analysis, but this would have seriously undermined the generalizability of any findings. A range of clinical and demographic variables was included that was thought to be useful predictors. It has been suggested that an episodic history of spine pain and intermittent symptoms affected by mechanical loading factors could be identifiers of responders to the McKenzie method¹⁷. These items were not included. A larger sample size would be needed to test out larger numbers of variables; for instance, to include 20 baseline variables would require at least 200 cases²⁸.

Previous studies that have sought to identify clinical prediction rules have been more successful at establishing clinical features associated with responders to manipulation and stabilization exercises¹²⁻¹⁴. These studies tended

TABLE 3. Predicted probability of success from model using liberal definition.

Duration	Region	Pain Change Status	Predicted Probability
Less than 12 weeks	Back	No centralization or abolition	0.447
Less than 12 weeks	Back	Centralization or abolition	0.689
Less than 12 weeks	Neck	No centralization or abolition	0.182
Less than 12 weeks	Neck	Centralization or abolition	0.379
12 weeks or more	Back	No centralization or abolition	0.100
12 weeks or more	Back	Centralization or abolition	0.234
12 weeks or more	Neck	No centralization or abolition	0.030
12 weeks or more	Neck	Centralization or abolition	0.077

TABLE 4. Predicted probability of success from model using strict definition.

Duration	Region	Predicted Probability
Less than 12 weeks	Back	0.491
Less than 12 weeks	Neck	0.130
12 weeks or more	Back	0.084
12 weeks or more	Neck	0.014

to look at a much more homogeneous population, which was acute to sub-acute back pain; they also included a very large number of baseline variables for univariate analysis. For instance, in the development of a clinical prediction rule for identifying success with manipulation, Flynn et al¹³ included about 50 items for initial analysis with only 71 patients. Of the 5 items included in the final model, it is our impression that two items, hip medial rotation greater than 35° and lower Fear Avoidance Belief Questionnaire scores, were somewhat counter-intuitive to the clinical relationship with response to manipulation. Another variable, symptom duration less than 16 days, seriously affected the clinical utility of the rule as acute patients are uncommonly seen in traditional physiotherapy settings. It might be suggested that initial inclusion of a large number of potential predictors is more likely to dredge up some significant predictive variables that might be of spurious clinical worth³⁵.

Age and gender of patients did not predict outcome, nor did QTF category, i.e., the extent of referred pain. This latter point is potentially of interest as it conflicts with previous findings that have associated leg pain with poorer outcomes⁴⁰⁻⁴². Also the addition of therapist overpressures or mobilization did not predict success more than patient-generated force only.

Clinical Impression

The clinical implications from this study suggest that treatment after three months has a limited probability of success, and that back pain rather than neck pain is treated more successfully. It also lends further support to the prognostic validity of centralization, at least short-term. However as only the latter factor can actually be influenced by therapists, this is probably the most important clinical predictor to come from the study. Our findings would suggest that patients with referred symptoms did as well as those with back or neck pain only. Equally, our findings would suggest that patients managed with exercise only did as well as those who also received mobilization.

A number of research implications can be drawn from this study. We have been relatively unsuccessful at finding useful predictive factors to explain a large proportion of success; this may be because our definition of success was too ambitious for the population or we failed to include other relevant factors at baseline. Nonetheless, we think the attempt to define sub-groups of the spine pain population who respond to specific treatments is a positive direction for future research. This may take the form of secondary analysis, such as the present work, the development of clinical prediction rules¹²⁻¹⁴, or trials in which classification-based treatment is compared to other management strategies^{10,15,16}. Work to date highlights several relevant issues in the first two of these study designs: the definition of treatment success and the included factors requires careful consideration. Inclusion of too many variables is also a concern. At present, the latter two study designs have been more successful at identifying useful clinical sub-groups.

Conclusion

Following a randomized controlled trial in which there were few significant differences between the McKenzie method and a cognitive behavioral approach, a secondary analysis was conducted to identify any clinical characteristics that distinguished those patients who responded best to the McKenzie method. According to our definitions of success, treatment successes were limited in number. Back pain (rather than neck pain), pain duration, and centralization response all predicted success, but the latter was not included in the strict definition. All variables were retained in the multiple regression, and when all 3 factors were included, the predicted probability of success was 69% under the liberal definition.

REFERENCES

- Hillman M, Wright A, Rajaratnam G, Tennant A, Chamberlain MA. Prevalence of low back pain in the community: Implications for service provision in Bradford, UK. *J Epidemiol Comm Health* 1996;50:347-352.
- Linton SJ, Hellsing AL, Hallden K. A population-based study of spinal pain among 35-45-year-old individuals. *Spine* 1998;23:1457-1463.
- Waxman R, Tennant A, Helliwell P. A prospective follow-up study of low back pain in the community. *Spine* 2000;25:2085-2090.
- Leclerc A, Niedhammer I, Landre MF, Ozguler A, Eto P, Pietri-Taleb F. One-year predictive factors for various aspects of neck disorders. *Spine* 1999;24:1455-1462.
- Picavet H, Schouten JSAG. Musculoskeletal pain in the Netherlands: Prevalence, consequences and risk groups—the DMC3-study. *Pain* 2003;102:167-178.
- Hill J, Lewis M, Papageorgiou AC, Dziedzic K, Croft P. Predicting persistent neck pain: A 1-year follow-up of a population cohort. *Spine* 2004;29:1648-1654.
- Deyo RA. Treatments for back pain: Can we get past trivial effects? *Ann Intern Med* 2004;141:957-958.
- Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. *BMJ* 2006;332:1430-1434.
- Bouter LM, van Tulder MW, Koes BW. Methodological issues in low back pain research in primary care. *Spine* 1998;23:2014-2020.
- Long A, Donelson R, Fung T. Does it matter which exercise? A randomized controlled trial of exercise for low back pain. *Spine* 2004;29:2593-2602.
- Cook C, Hegedus EJ, Ramey K. Physical therapy exercise intervention based on classification using the patient response method: A systematic review of the literature. *J Man Manipulative Ther* 2005;13:152-162.
- Childs JD, Fritz JM, Flynn TW, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: A validation study. *Ann Intern Med* 2004;141:920-928.
- Flynn T, Fritz J, Whitman J, Wainner R, Mangel J, Rendeiro D. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine* 2002;27:2835-2843.
- Hicks GE, Fritz JM, Delitto A, McGill SM. Preliminary development of a clinical prediction rule for determining which patients with low back pain will respond to a stabilization exercise program. *Arch Phys Med Rehabil* 2005;86:1753-1762.
- Fritz JM, Delitto A, Erhard RE. Comparison of classification-based physical therapy with

- therapy based on clinical practice guidelines for patients with acute low back pain: A randomized controlled trial. *Spine* 2003;28:1363–1372.
16. Brennan GP, Fritz JM, Hunter SJ, Thackeray A, Delitto A, Erhard RE. Identifying subgroups of patients with acute/sub=acute “non-specific” low back pain. *Spine* 2006;31:623–631.
17. McKenzie RA. *The Lumbar Spine: Mechanical Diagnosis and Therapy*. Waikanae, New Zealand: Spinal Publication Ltd, 1981.
18. McKenzie RA. *The Cervical and Thoracic Spine: Mechanical Diagnosis and Therapy*. Waikanae, New Zealand: Spinal Publication Ltd, 1990.
19. McKenzie RA, May S. *The Lumbar Spine Mechanical Diagnosis and Therapy*. 2nd ed. Waikanae, New Zealand: Spinal Publication New Zealand Ltd, 2003.
20. McKenzie RA, May S. *The Cervical and Thoracic Spine Mechanical Diagnosis and Therapy*. 2nd ed. Waikanae, New Zealand: Spinal Publication New Zealand Ltd, 2006.
21. Foster NE, Thompson KA, Baxter GD, Allen JM. Management of nonspecific low back pain by physiotherapists in Britain and Ireland: A descriptive questionnaire of current clinical practice. *Spine* 1999;24:1332–1342.
22. Gracey JH, McDonough SM, Baxter GD. Physiotherapy management of low back pain: A survey of current practice in Northern Ireland. *Spine* 2002;27:406–411.
23. Jackson DA. How is low back pain managed? Retrospective study of the first 200 patients with low back pain referred to a newly established community-based physiotherapy department. *Physiotherapy* 2001;87:573–581.
24. Aina A, May S, Clare H. The centralization phenomenon of spinal symptoms: A systematic review. *Man Ther* 2004;9:134–143.
25. Clare HA, Adams R, Mahler CG. A systematic review of efficacy of McKenzie therapy for spinal pain. *Aust J Physiother* 2004;50:209–216.
26. Machado LAC, de Souza MvS, Ferreira PH, Ferreira ML. The McKenzie method for low back pain: A systematic review of the literature with a meta-analysis approach. *Spine* 2006;31:E254–E262.
27. Klaber-Moffett J, Jackson DA, Gardiner ED, et al. Comparison of “McKenzie” physical therapy treatment vs. a brief physiotherapy intervention: A randomized clinical trial of neck pain and back patients in primary care. *Rheumatology* 2006;45:1514–1521.
28. Haynes RB, Sackett DL, Guyatt GH, Tugwell P. *Clinical Epidemiology: How to Do Clinical Practice Research*. 3rd ed. Philadelphia: Lippincott Williams & Wilkins, 2006.
29. Harrell FE. *Regression Modelling Strategies: With Applications to Linear Models, Logistic Regression and Survival Analysis*. New York: Springer Verlag, 2002.
30. Clare HA, Adams R, Maher CG. Reliability of McKenzie classification of patients with cervical or lumbar pain. *J Manip Physiol Ther* 2005;28:122–127.
31. Kilpikoski S, Airaksinen O, Kankaanpää M, Leminen P, Videman T, Alen M. Inter-tester reliability of low back pain assessment using the McKenzie method. *Spine* 2002;27:E207–E214.
32. Razmjou H, Kramer JF, Yamada R. Inter-tester reliability of the McKenzie evaluation in assessing patients with mechanical low-back pain. *J Orthop Sports Phys Ther* 2000;30:368–389.
33. Fritz JM, Delitto A, Vignovic M, Busse RG. Interrater reliability of judgements of the centralization phenomenon and status change during movement testing in patients with low back pain. *Arch Phys Med Rehabil* 2000;81:57–61.
34. Werneke M, Hart DL, Cook D. A descriptive study of the centralization phenomenon. *Spine* 1999;24:676–683.
35. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules: A review and suggested modifications of methodological standards. *JAMA* 1997;277:488–494.
36. Jellema P, van der Horst H, Vlaeyen JWS, Stalman WAB, Bouter LM, van der Windt DAWM. Predictors of outcome in patients with (sub)acute low back pain differ across treatment groups. *Spine* 2006;31:1699–1705.
37. Stratford PW, Binkley J, Solomon P, Finch E, Gill C, Moreland J. Defining the minimum level of detectable change for the Roland-Morris Questionnaire. *Phys Ther* 1996;76:359–365.
38. Pool JJM, Ostelo RWJG, Hoving JL, Bouter LM, de Vet HCW. Minimally clinically important change of the neck disability index and the numerical rating scale for patients with neck pain. *Spine* 2007;32:3047–3051.
39. Cherkin DC, Deyo RA, Street JH, Barlow W. Predicting poor outcomes for back pain seen in primary care using patients’ own criteria. *Spine* 1996;21:2900–2907.
40. Goertz MN. Prognostic indicators for acute low back pain. *Spine* 1990;15:1307–1310.
41. Thomas E, Silman AJ, Croft PR, Papageorgiou AC, Jayson MI, MacFarlane GJ. Predicting who develops chronic low back pain in primary care: A prospective study. *BMJ* 1999;318:1662–1667.