

ORIGINAL ARTICLE

Preliminary Development of a Clinical Prediction Rule for Determining Which Patients With Low Back Pain Will Respond to a Stabilization Exercise Program

Gregory E. Hicks, PhD, PT, Julie M. Fritz, PhD, PT, ATC, Anthony Delitto, PhD, PT, Stuart M. McGill, PhD

ABSTRACT. 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; 86:1753-62.

Objective: To develop a clinical prediction rule to predict treatment response to a stabilization exercise program for patients with low back pain (LBP).

Design: A prospective, cohort study of patients with nonradicular LBP referred to physical therapy (PT).

Setting: Outpatient PT clinics.

Participants: Fifty-four patients with nonradicular LBP.

Intervention: A standardized stabilization exercise program.

Main Outcome Measure: Treatment response (success or failure) was categorized based on changes in the Oswestry Disability Questionnaire scores after 8 weeks.

Results: Eighteen subjects were categorized as treatment successes, 15 as treatment failures, and 21 as somewhat improved. After using regression analyses to determine the association between standardized examination variables and treatment response status, preliminary clinical prediction rules were developed for predicting success (positive likelihood ratio [LR], 4.0) and failure (negative LR, .18). The most important variables were age, straight-leg raise, prone instability test, aberrant motions, lumbar hypermobility, and fear-avoidance beliefs.

Conclusions: It appears that the response to a stabilization exercise program in patients with LBP can be predicted from variables collected from the clinical examination. The prediction rules could be used to determine whether patients with LBP are likely to benefit from stabilization exercises.

Key Words: Classification; Exercise; Low back pain; Rehabilitation.

© 2005 by the American Congress of Rehabilitation Medicine and the American Academy of Physical Medicine and Rehabilitation

From the Departments of Physical Therapy and Rehabilitation Science and Epidemiology and Preventive Medicine, University of Maryland School of Medicine, Baltimore, MD (Hicks); Division of Physical Therapy, University of Utah, Salt Lake City, UT (Fritz); Department of Physical Therapy, University of Pittsburgh, Pittsburgh, PA (Delitto); and Department of Kinesiology, University of Waterloo, Waterloo, ON, Canada (McGill).

Supported by the Foundation for Physical Therapy Clinical Research Center.

The opinions and assertions contained herein are the private views of the author and are not to be construed as official or as reflecting the views of the U.S. Department of the Army or the Department of Defense.

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit upon the author(s) or upon any organization with which the author(s) is/are associated.

Correspondence to Gregory E. Hicks, PhD, PT, Dept of Physical Therapy and Rehabilitation Science, University of Maryland School of Medicine, 100 Penn St, Baltimore, MD 21201, e-mail: ghicks@som.umaryland.edu. Reprints are not available from the author.

0003-9993/05/8609-9715\$30.00/0

doi:10.1016/j.apmr.2005.03.033

A PATHOLOGY-BASED DIAGNOSTIC approach to low back pain (LBP) sets forth the idea that LBP is defined by a deviation of the lumbopelvic complex from its normative physiologic and anatomic state.¹ On application of the pathology-based model to LBP, one must first identify a structural fault and then treatment can be directed at correction of the fault. As a result of this treatment and subsequent correction, the signs and symptoms should dissipate.² A pathology-based approach to diagnosis for LBP has proven difficult because of the inability to identify a structural pathology in the vast majority of patients with LBP.³

Despite the difficulties in identifying structural pathology for most patients, many authors continue to suggest that patients with LBP are not a homogenous group, but rather should be classified into subgroups that share similar clinical characteristics (eg, age, symptom duration, distribution).⁴⁻⁶ This type of classification system could guide diagnosis and treatment and improve overall decision making in the management of LBP patients.⁷ It may also improve research by denoting homogeneous subgroups for treatment outcome studies.⁸

Patients with lumbar segmental instability (LSI) have been proposed as a unique subgroup of patients with LBP.^{7,9-11} LSI has been defined as a condition in which there is a loss of stiffness between spinal motion segments, such that normally tolerated external loads result in pain or deformity or place neurologic structures at risk.¹⁰ The diagnosis of LSI has traditionally relied on the use of lateral flexion and extension radiographs.¹²⁻¹⁴ Thresholds for defining LSI have been developed from studies that have examined the magnitude of rotational and translational movement normally available at each spinal motion segment during flexion and extension of the lumbar spine.^{14,15} However, the use of these criteria for defining LSI has proven unsatisfactory because of high false-positive rates.^{16,17} For example, Hayes et al¹⁷ found that 42% of asymptomatic subjects had at least 1 segment exceeding the instability thresholds. The inadequacies associated with the use of imaging to diagnosis LSI have led to a greater emphasis on clinical characteristics that may be indicative of the condition. Historical variables, presence of an "instability catch," or excessive intervertebral mobility have been suggested as diagnostic¹⁸⁻²⁰; however, the validity of these findings is mostly unknown. Because of the current lack of a true reference standard for confirmation of LSI, it is not possible to accurately diagnose LSI clinically; therefore, another approach should be considered.

If LSI could be accurately diagnosed, the conservative treatment of choice would be a lumbar stabilization program to address the tissue damage and resultant loss of spinal stiffness caused by mechanical overload to the spinal stabilizing system. The goals of stabilization exercise are to train muscular motor patterns to increase spinal stability, restrain aberrant micromotion, and reduce associated pain. Although several studies that observed muscle onset and electromyographic patterns have suggested that certain muscles are important stabilizers,²¹ measurement of stability was not performed. In contrast, those

studies that have quantified stability^{22,23} unanimously agree that all muscles play a role in ensuring spine stability and that the motor patterns of cocontraction between the full complement of muscles are of utmost importance to ensure stability and minimize pain. Although the exercises used in this study are often described in the literature as targeting a specific muscle group, electromyography studies show that the exercises will challenge groups of muscles in a way that enhances overall spinal stability.

Recent research²⁴⁻²⁶ has shown that exercise programs can be designed to challenge the stabilizing muscles of the spine in a way that appears to be effective for at least certain subgroups of patients with LBP. The development of a valid method for identifying this subgroup of patients would help clinicians in determining which patients with LBP are most likely to benefit from this exercise program. The purpose of this study was to explore the value of various demographic, historic, and clinical examination variables for predicting outcome after participation in a program of spinal stabilization exercises. The reference standard used to determine success was change in disability.

METHODS

We conducted a prospective cohort study involving patients referred to physical therapy (PT) with LBP. Subjects were recruited from new patients referred for treatment of LBP to 1 of 3 outpatient PT clinics located in the Pittsburgh, PA, area or the outpatient PT clinic at Keesler Air Force Base Medical Center (Biloxi, MS). All new patients with complaints of LBP, with or without leg pain, over the age of 18 were eligible for participation. Exclusion criteria were as follows: previous spinal fusion surgery; LBP attributable to current pregnancy; acute fracture, tumor, or infection; and presence of 2 or more of the following signs of nerve root compression: diminished lower-extremity strength, sensation, or reflexes. All subjects reviewed and signed a consent form approved by the University of Pittsburgh Health Sciences Institutional Review Board.

Therapists

Four licensed physical therapists participated in the examination and treatment of subjects in this study. All therapists received written instructions and specific training regarding all study procedures. Each therapist had to pass a written test to show an acceptable level of knowledge regarding subject recruitment, clinical examination, and exercise progression.

Examination Procedures

At baseline, subjects were evaluated by the treating physical therapist. The evaluation consisted of the collection of basic demographic information, subject self-report measures, questions of history, and physical examination measures.

Each subject completed a numeric pain rating to assess severity of pain by using an 11-point (0–10) scale.²⁷ The Fear Avoidance Beliefs Questionnaire (FABQ)²⁸ was used to measure the level of fear-avoidance beliefs. The FABQ has 2 subscales that measure fear-avoidance beliefs about work (7-item scale) and physical activity (4-item scale). The Modified Oswestry Low Back Pain Disability Questionnaire (ODQ)²⁹ is a disease-specific measure of disability among patients with LBP. This questionnaire has been used extensively in randomized trials and has shown excellent reliability and good construct validity in comparison with other pain and disability measures.²⁹⁻³¹ The ODQ was assessed at baseline and after 8 weeks of treatment and served as the reference standard for determining the success of the treatment program.

All subjects underwent a standardized history and physical examination. History taking included questions proposed to indicate the presence of LSI,^{7,18,19} including mode of onset, duration of symptoms, number of previous episodes of LBP, and response to various interventions for previous episodes. Subjects were also asked about the distribution of symptoms for the current episode and were asked to rank sitting versus standing and walking and morning versus evening with respect to their symptoms.

Operational definitions for components of the physical examination are included in appendix 1. The physical examination included lumbar spine range of motion (ROM) measured with a single inclinometer. Waddell et al³² found these measurements to be highly reliable (intraclass correlation coefficient [ICC] range, .86–.95). Any aberrant motions believed to be associated with LSI occurring during the performance of lumbar ROM were noted, including an instability catch,^{33,34} painful arc of motion,³⁵ “thigh climbing” (Gower’s sign), or a reversal of lumbopelvic rhythm.⁷ Subjects were categorized as either having or not having aberrant motions present, a determination found to be reliable in a previous study.³⁶ Posterior-anterior lumbar mobility testing³⁷ was performed at each spinal level. Although the reliability coefficients for these judgments generally have been shown to be only fair,^{38,39} percentage agreement among examiners is high³⁶ and mobility judgments have been shown to be useful for determining which patients with LBP may respond to a manipulation intervention.⁴⁰ Generalized ligamentous laxity was assessed on a 9-point scale described by Beighton and Horan.⁴¹ Higher numbers indicate greater laxity. We previously found the reliability of this scale to be excellent (ICC=.79).³⁶ Two special tests proposed to indicate LSI, the prone instability⁴² and posterior shear⁷ tests, were performed. We found the prone instability test ($\kappa=.87$) to be more reliable than the posterior shear test ($\kappa=.35$),³⁶ but the validity of these tests has not been studied previously.

Muscle endurance of the spinal extensors and lateral flexors was assessed by using the modified Biering-Sorenson extensor endurance test and the side support test, respectively, which have been found to be highly reliable (ICC range, .93–.99).⁴³ Two strength tests, the active sit-up and active straight-leg raising (SLR) tests, were also used. These tests have been found to be both reliable (κ range, .48–.77) and valid in a previous study.³²

Treatment

All study participants underwent PT treatment consisting of the same stabilization program. Subjects were scheduled to attend supervised therapy sessions twice weekly for an 8-week period and were asked to perform exercises at home daily. Subjects were asked to complete a compliance log for home exercises to verify performance. The treatment program used was a specific stabilization program based on current evidence from the biomechanical and electromyography studies.^{25,44-46} The exercise program was designed to challenge and encourage stabilizing motor patterns for the primary stabilizing muscles of the spine including the rectus abdominus, transversus abdominus,^{44,45} internal oblique abdominals,⁴⁷ erector spinae and multifidus,^{25,44,48,49} and quadratus lumborum.⁵⁰ The stabilization exercise program focused on encouraging repeated submaximal efforts to mimic the function of these muscles in spine stabilization.^{44,46} The exercises for each muscle group and criteria for progression are listed in table 1. Progression of the exercise program was directed by the physical therapist based on the criteria given.

Table 1: Stabilization Exercises With Criteria for Progression of Each Exercise

Primary Muscle Group*	Exercises	Criteria for Progression
Transversus abdominus	Abdominal bracing	30 repetitions with 8-s hold
	Bracing with heel slides	20 repetitions per leg with 4-s hold
	Bracing with leg lifts	20 repetitions per leg with 4-s hold
	Bracing with bridging	30 repetitions with 8-s hold, then progress to 1 leg
	Bracing in standing	30 repetitions with 8-s hold
	Bracing with standing row exercise	20 repetitions per side with 6-s hold
Erector spinae/multifidus	Quadruped arm lifts with bracing	30 repetitions with 8-s hold on each side
	Quadruped leg lifts with bracing	30 repetitions with 8-s hold on each side
	Quadruped alternate arm and leg lifts with bracing	30 repetitions with 8-s hold on each side
Quadratus lumborum	Side support with knees flexed	30 repetitions with 8-s hold on each side
	Side support with knees extended	30 repetitions with 8-s hold on each side
Oblique abdominals	Side support with knees flexed	30 repetitions with 8-s hold on each side
	Side support with knees extended	30 repetitions with 8-s hold on each side

*Although certain muscle groups are preferentially activated with each exercise sequence, each exercise progression will promote stability by producing motor patterns of cocontraction among all spinal stabilizing muscles.

Data Analysis

The outcome of treatment was determined based on the changes in the ODQ between the baseline and 8-week assessments. For each subject, the points of change in the ODQ score over 8 weeks (initial ODQ score–8-wk ODQ score), and the percentage change in the ODQ score ($[(\text{initial ODQ score} - 8\text{-wk ODQ score}) / \text{initial ODQ score}] \times 100\%$) were calculated. Any subject with a percentage change of 50% or greater was classified as a stabilization treatment success. For subjects with less than a 50% improvement, the points of change in the ODQ score were examined. Subjects experiencing less than 6 points of improvement on the ODQ were classified as stabilization treatment failures. All other subjects with at least 6 points of improvement but less than 50% improvement were classified as improved with stabilization treatment.

The rationale for the criteria for defining stabilization success and failure comes from the literature and previous experience with the ODQ. In previous studies, patients with LBP believed to be appropriately matched to their treatment program experienced an improvement in ODQ scores from 57% to 83%, whereas patients receiving unmatched treatments experienced improvements ranging from 20% to 38% over a 1- to 4-week treatment period.^{7,51-54} In addition, O'Sullivan et al²⁶ used a similar exercise program for patients with spondylolysis

or spondylolysis and reported a 48% improvement after 10 weeks in the intervention group. Based on these findings, we believe that subjects experiencing at least a 50% improvement in ODQ score over 8 weeks responded successfully to the stabilization intervention, and clinical decision making would be enhanced if clinicians had a mechanism for identifying these individuals. The minimum clinically important difference (MCID) in ODQ score has been calculated as 5 to 6 points.^{7,29,54,55} Given the generally favorable natural history of LBP, we believe that subjects who did not even achieve the MCID in ODQ score over 8 weeks of treatment were appropriately considered to have failed the intervention. The ability to identify such patients a priori would allow clinicians to consider alternative interventions.

Individual variables from subjects' demographic, self-report, history, and physical examination were tested for a significant univariate association with stabilization success by using independent sample *t* tests for continuous variables and chi-square tests for categorical variables. Univariate significance of individual variables with stabilization failure was also assessed. Variables with a significant univariate association with either success or failure ($P < .10$) were retained as potential predictor variables for that outcome. A more liberal significance level was used because this step was intended to filter variables with no association with outcome, and we did not wish to exclude

Table 2: History and Demographic Variables Assessed at Baseline

History/Demographic Variables	All Subjects (N=54)	Succeeded With Treatment (n=18)	Improved With Treatment (n=21)	Failed With Treatment (n=15)
Age (y)	42.4±12.7	38.2±13.4	43.2±11.8	46.5±12.5
Sex (% women)	57.4	55.6	61.9	53.3
Mode of onset (% sudden)	33.3	33.3	33.3	33.3
Duration of current symptoms (d)	40.6±44.2	35.3±41.5	35.7±38.0	52.0±54.0
Distribution of symptoms				
Back/buttock only (%)	53.7	66.7	47.6	46.7
Distal symptoms (%)	46.3	33.3	52.4	53.3
Prior history of LBP (%)	70.4	72.2	71.4	66.7
More than 3 prior episodes	59.3	72.2	61.9	40.0
Episodes of LBP becoming more frequent (%)	40.1	44.4	52.4	20.0
Improved with prior bracing treatment (%)	3.7	5.6	4.8	0
Improved with prior manipulation treatment (%)	25.9	16.7	28.6	33.3

NOTE. Values are mean ± standard deviation (SD) unless otherwise indicated.

Table 3: Self-Report Variables

Self-Report Variable	All Subjects (N=54)	Succeeded With Treatment (n=18)	Improved With Treatment (n=21)	Failed With Treatment (n=15)
Pain rating	4.5±2.4	4.3±2.4	5.2±2.6	3.3±1.9
FABQ work subscale	13.9±12.0	13.6±10.5	17.2±12.3	9.5±12.4
FABQ physical activity subscale	14.6±5.9	15.1±6.4	16.2±4.9	11.9±5.9
Baseline ODQ score	29.7±13.7	25.3±11.4	39.3±11.5	21.6±11.1
Change in ODQ score with treatment (points)	10.0±10.3	18.1±7.3	11.3±4.8	-2.4±7.4
Change in ODQ score with treatment (%)	34.6±39.3	74.8±18.5	30.0±10.7	-10.0±33.7

NOTE. Values are mean ± SD.

any potentially useful predictor variables. Any continuous variable found to have a significant univariate association with outcome was further explored by constructing a receiver operator characteristic (ROC) curve. The ROC curve was used to determine a cutoff point defining a positive test.^{7,54,56} The point nearest the upper left hand corner of an ROC curve represents the cutoff score with the best diagnostic accuracy.⁵⁷ This point^{7,54} was therefore selected as the cutoff score defining a positive test. Accuracy statistics (sensitivity, specificity, positive and negative likelihood ratio [LR] values) with the corresponding 95% confidence interval (CI) were calculated based on this cutoff score. For categorical variables with significant univariate relationships, accuracy statistics with 95% CI were calculated.⁵⁸

For the prediction of success, the "target condition" used to calculate accuracy statistics was success with stabilization. The focus for the prediction of success was on maximizing the positive LR because this value represents the change in odds favoring success given a positive diagnostic test result. For the prediction of failure, the "target condition" was not failing with stabilization (ie, experiencing either improvement or success). The focus for the prediction of failure was therefore on minimizing the negative LR value. A test with a small negative LR would serve as a useful screening tool because this indicates that a negative result would substantially decrease the odds of a favorable result (improvement or success) with stabilization treatment.

All variables with a significant univariate relation to treatment outcome were then used to develop multivariate clinical prediction rules. If more than 5 variables were found to have a significant relation to an outcome, then the variables were entered into a forward stepwise logistic regression equation to reduce the number of predictors. A significance of .15 was required to enter, and a significance of .20 was used to remove a variable from the equation. More liberal significance levels were used because this is the first approximation of a clinical prediction rule. Once the number of predictor variables was determined, the clinical prediction rule was developed by examining the accuracy statistics for various combinations of the retained variables.

RESULTS

A total of 57 subjects gave informed consent for participation between September 2001 and May 2002. Three subjects initially agreed to participate but withdrew after several days because they did not want to commit to the time required. Because no data were obtained from these subjects, they were not included in the analysis. One subject began the stabilization program but dropped out after 2 weeks for an unknown reason. This subject was included in the analysis and considered a stabilization failure. A total of 54 subjects were therefore included in the analysis.

Table 2 provides descriptive information. The mean age ± standard deviation (SD) was 42.4±12.7 years; 57.4% were women and 85.2% were white. Thirty-eight (70.4%) subjects reported a history of LBP that required activity modification. The mean ODQ score at baseline was 29.7±13.7 (table 3). Eighteen (33.3%) subjects experienced a 50% or greater improvement in ODQ score over 8 weeks and were categorized as stabilization successes. Of the remaining 36 subjects, 15 (27.8%) experienced less than a 6-point improvement on the ODQ and were categorized as stabilization failures. The remaining 21 (38.9%) subjects were categorized as improved with stabilization treatment. Initial and follow-up ODQ scores for the 3 outcome groups are shown in figure 1. The mean number of PT visits attended during the 8-week treatment period was 9.7±2.2, and there was no association between the number of visits and change in ODQ score over the 8-week period ($P>.05$).

Tables 3 and 4 show the results of the history and demographic and physical examination findings. Four variables were significantly related to success and were retained as potential predictors: age (cutoff value defining a positive test <40y), average SLR (cutoff value defining a positive test >91°), the presence of aberrant movement during lumbar ROM, and a positive prone instability test (table 5). Nine variables were retained as potential predictors of failure (table 6), including the FABQ physical activity subscale (cutoff value defining a positive test >8 points), pain rating (cutoff value defining a positive test >2), lumbar flexion ROM (cutoff value defining a positive test >33°), discrepancy in SLR (cutoff value defining a positive test <10°), 3 or more previous episodes of LBP, increasing frequency of episodes of LBP, aberrant movement during lumbar ROM, the presence of hypermobility during lumbar spring testing, and a positive prone instability test.

Among the individual predictor variables for success, age less than 40 years had the greatest positive LR (3.7; 95% CI,

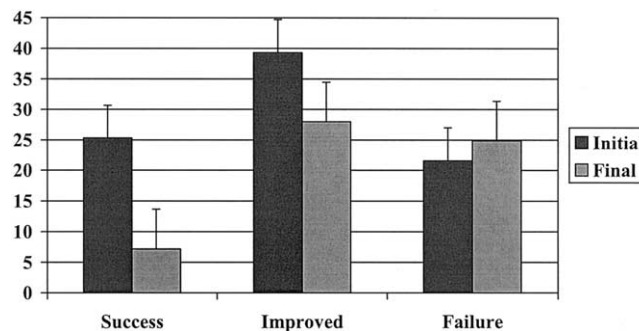


Fig 1. Initial and final ODQ scores for the 3 outcome groups.

Table 4: Physical Examination Variables Assessed at Baseline

Physical Examination Variables	All Subjects (N=54)	Succeeded With Treatment (n=18)	Improved With Treatment (n=21)	Failed With Treatment (n=15)
Total flexion (deg)	88.5±22.9	94.2±25.8	84.2±25.9	87.5±12.3
Pelvic flexion (deg)	47.3±18.4	49.9±23.1	42.2±17.9	51.2±10.6
Lumbar flexion (deg)	41.5±13.6	44.3±12.4	42.3±15.7	36.9±11.1
Percentage of total flexion from lumbar spine (deg)	47.8±14.0	49.2±14.3	50.8±15.2	42.0±10.6
Total extension (deg)	25.9±9.9	26.3±7.5	23.8±10.4	28.3±11.6
Left side-bending (deg)	27.6±8.4	28.3±6.7	28.0±9.3	26.2±9.2
Right side-bending (deg)	27.9±9.4	27.2±9.8	28.9±10.0	27.3±8.5
Average side-bending (deg)	28.0±8.1	28.5±6.7	28.4±9.2	26.7±8.3
Side-bending discrepancy* (deg)	4.7±5.2	5.4±7.3	3.9±3.8	4.8±3.8
Left SLR (deg)	77.6±13.7	82.4±11.0	75.4±15.6	74.7±13.0
Right SLR (deg)	77.9±14.6	82.6±13.9	76.4±14.4	74.2±15.2
Average SLR (deg)	77.7±13.7	82.5±12.1	75.9±14.7	74.5±13.5
SLR discrepancy* (deg)	5.6±4.3	5.2±4.3	4.9±4.4	7.2±3.9
Aberrant motion during lumbar ROM (% yes)	59.3	77.8	66.7	26.7
Beighton scale score	0.9±1.7	1.1±1.9	0.9±1.4	0.6±2.1
Posterior shear test (% positive)	42.6	33.3	57.1	33.3
Prone instability test (% positive)	51.9	72.2	61.9	13.3
Active sit-up test (% positive)	13.0	11.1	14.3	11.1
Active SLR test (% positive)	20.3	16.7	28.6	13.3
Any hypermobility present during spring testing (% yes)	20.3	22.2	33.3	0
Side support left (s)	27.7±24.9	32.1±23.2	27.1±26.8	23.3±25.1
Side support right (s)	26.3±24.6	29.7±21.9	25.9±27.0	22.7±25.4
Average side support (s)	27.0±24.5	30.9±22.2	26.5±26.5	23.0±25.0
Side support discrepancy* (s)	6.2±5.5	5.6±6.1	7.1±5.8	5.6±4.3
Extensor endurance (s)	49.0±45.8	51.7±38.5	48.3±43.7	46.7±58.4
Ratio average side support/ extensor endurance (%)	94.1±109.7	74.7±47.6	102.2±133.4	108.8±135.3

NOTE. Values are mean ± SD unless otherwise indicated.

*Discrepancy values calculated as the absolute value of the right side minus the left side.

1.6–8.3). This finding indicates that a subject under the age of 40 had a 3.7 higher odds of succeeding with stabilization treatment. For the prediction of failure, the best individual screening test was a FABQ physical activity subscale score of greater than 8 (negative LR, .26; 95% CI, .08–.78), indicating that a negative result (ie, a score of ≤8) would decrease the odds of experiencing some improvement with stabilization treatment by .26.

The 4 predictors of success with stabilization were considered for the multivariate clinical prediction rule. The best rule for predicting success was the presence of 3 or more of the 4 variables (positive LR, 4.0; 95% CI, 1.6–10.0) (table 7). The 9 potential predictor variables for failure were entered into a stepwise logistic regression equation. Four variables were retained in the final model (table 8) (FABQ physical activity subscale score >8, aberrant movement, prone instability test,

hypermobility during lumbar spring testing). The best rule for predicting failure occurred when a positive test was defined as the presence of 2 or more of the 4 variables (negative LR, .18; 95% CI, .08–.38; positive LR, 6.3; 95% CI, 1.7–23.2) (see table 8). This means that when only 1 or none of these 4 variables was present, the odds of experiencing some improvement with stabilization treatment decreased by a factor of .18. If 2 or more of the 4 variables were present, the odds of some improvement with stabilization increased by 6.3.

DISCUSSION

Exercises designed to improve spinal stabilization have gained popularity in the conservative treatment of patients with LBP; however, the evidence for the effectiveness of this approach is sparse and equivocal.⁵⁹ A recent systematic review⁶⁰ of exercise therapy for LBP concluded that exercise therapy

Table 5: Significant Univariate Predictors of Stabilization Success

Variable Associated With Success	Signif	Sensitivity	Specificity	Positive LR	Negative LR
Age (<40y)	.082	.61 (.39–.80)	.83 (.68–.92)	3.7 (1.6–8.3)	.47 (.26–.85)
Average SLR (>91°)	.069	.28 (.13–.51)	.92 (.78–.97)	3.3 (.90–12.4)	.79 (.58–1.1)
Aberrant movement present	.050	.78 (.55–.91)	.50 (.35–.66)	1.6 (1.0–2.3)	.44 (.18–1.1)
Positive prone instability test	.034	.72 (.49–.88)	.58 (.42–.73)	1.7 (1.1–2.8)	.48 (.22–1.1)

NOTE. Values are accuracy statistics (95% CIs).

Abbreviation: Signif, significance.

Table 6: Significant Univariate Predictors of Stabilization Failure

Variable Associated With Failure	Signif	Sensitivity	Specificity	Positive LR	Negative LR
FABQ physical activity subscale (<9)	.032	.90 (.76-.96)	.40 (.20-.64)	1.4 (.98-2.3)	.26 (.08-.78)
Initial pain rating (<3)	.042	.77 (.62-.87)	.40 (.20-.64)	1.3 (.82-2.0)	.58 (.25-1.3)
Discrepancy in SLR (>10°)	.093	.90 (.76-.96)	.40 (.20-.64)	1.4 (.95-2.2)	.32 (.11-.90)
Percentage of total flexion from the lumbar spine (<37%)	.058	.80 (.65-.89)	.40 (.20-.64)	1.3 (.85-2.1)	.51 (.21-1.2)
Fewer than 3 prior LBP episodes	.074	.67 (.51-.79)	.60 (.36-.80)	2.0 (.81-4.7)	.76 (.30-1.0)
No increasing frequency of LBP episodes	.054	.49 (.34-.64)	.80 (.55-.93)	2.4 (.84-7.0)	.64 (.43-.95)
Aberrant movement absent	.003	.72 (.56-.84)	.73 (.48-.89)	2.7 (1.1-6.4)	.39 (.21-.69)
No lumbar hypermobility with spring testing	.021	.28 (.17-.44)	1.0 (.80-1.0)	9.2* (5.0-16.9)	.74* (.59-.96)
Negative prone instability test	.003	.67 (.51-.79)	.87 (.62-.96)	5.0 (1.4-18.5)	.39 (.24-.63)

NOTE. Values are accuracy statistics (95% CIs).

*LR values estimated by adding 0.5 to each cell to avoid division by zero error.

overall was not effective for patients with acute LBP but may be helpful for those with chronic LBP. The review also noted that the evidence could not be used to examine the effectiveness of specific types of exercise (eg, flexion, extension, strengthening, stabilization) because of heterogeneity in study populations and interventions.⁶⁰ The few studies that have examined specific stabilization exercise programs such as that used in this study in more homogenous populations have shown promising results. O'Sullivan et al²⁶ randomized patients with chronic, recurrent LBP who had a radiologic diagnosis of spondylolysis or spondylolisthesis to receive either stabilization exercises or usual care directed by a general practitioner and found statistically significant reductions in pain and disability at a 30-month follow-up. Hides et al^{24,25} studied patients with acute, first-time episode of unilateral LBP, comparing stabilization exercises particularly targeting the multifidus muscle with usual medical management. No significant differences in disability or pain were found after 4 weeks,²⁵ but the stabilization group experienced significantly fewer recurrences at 2- to 3-year follow-up.²⁴ These studies indicate that a specific stabilization exercise approach may be effective for certain subgroups of patients. This study is an initial step in identifying the subgroup of patients for whom this approach may be particularly effective.

The preliminary prediction rule for success with stabilization treatment contains 4 variables: positive prone instability test, aberrant movements present, average SLR greater than 91°, and age greater than 40 years old. These variables seem reasonable based on the available theoretical literature regarding LSI and stabilization. The prone instability test works on the premise that if pain is present on passive provocation testing of the vertebral levels but disappears when the spinal extensors are active, then the muscle activity may be able to stabilize the segment and reduce pain, and the individual may benefit from stabilization exercises. Several authors^{7,20,61} have suggested that the observation of aberrant movement patterns during

active trunk motion is valuable in the diagnosis of LSI. The presence of aberrant movements may represent an inability to adequately control lumbar motion and indicate a need for stabilization exercises. Decreased SLR ROM is related to the presence of radiculopathy and a generally worse prognosis.⁶²⁻⁶⁴ In our study, greater SLR mobility was associated with stabilization success. The final predictor of success was younger age (<40y). Increased age is associated with decreased lumbar extensor muscle mass.^{65,66} Despite the evidence suggesting detrimental effects to muscle with aging, there is encouraging evidence that older men and women can still increase muscle strength even up to age 90 with appropriate training.⁶⁷ Adults over the age of 39 may require more than 8 weeks of training to gain the same benefits as their younger peers with LBP. Additionally, because of the decreased flexibility and ROM associated with aging, it is also possible that older adults may require a combination of mobility-related interventions followed by stabilization exercise to achieve the same gains as their younger counterparts.

We also sought to develop a preliminary clinical prediction rule for patients likely to fail with stabilization treatment. The prediction rule contained 4 variables: negative prone instability test, aberrant movements absent, FABQ physical activity subscale score greater than 9, and no hypermobility with lumbar spring testing. A negative prone instability test may indicate that muscular stabilization is not needed, and the absence of an aberrant movement pattern during active trunk motion may indicate an ability to control movement in a pain-free fashion. The FABQ measures fear-avoidance beliefs in patients with LBP.²⁸ In this study, lower levels of fear-avoidance beliefs about physical activity were associated with stabilization failure. This is counter to other findings in the literature that suggest higher levels of fear-avoidance beliefs are associated with poorer outcomes.^{68,69} Our findings may suggest that this stabilization program preferentially helps patients with moderate to high levels of fear-avoidance beliefs to overcome their

Table 7: Clinical Prediction Rule for Predicting Success With Stabilization Treatment

No. of Variables Present	Sensitivity	Specificity	Positive LR	Negative LR
One or more	.94 (.74-.99)	.28 (.16-.44)	1.3 (1.0-1.6)	.20 (.03-1.4)
Two or more	.83 (.61-.94)	.56 (.40-.71)	1.9 (1.2-2.9)	.30 (.10-.88)
Three or more	.56 (.34-.75)	.86 (.71-.94)	4.0 (1.6-10.0)	.52 (.30-.88)

NOTE. Values are accuracy statistics (95% CIs). The best prediction rule based on the positive LR value is the presence of at least 3 of the predictor variables (positive prone instability test, aberrant movement present, average SLR >91°, age <40y).

Table 8: Clinical Prediction Rule for Predicting Failure With Stabilization Treatment

No. of Variables Present	Sensitivity	Specificity	Positive LR	Negative LR
One or more positive tests	.97 (.88–1.0)	.13 (.04–.38)	1.1 (.92–1.3)	.20 (.02–2.0)
Two or more positive tests	.85 (.70–.93)	.87 (.62–.96)	6.3 (1.7–23.2)	.18 (.08–.38)
Three or more positive tests	.59 (.43–.73)	1.0 (.80–1.0)	18.8* (10.9–32.3)	.43* (.29–.65)
Four or more positive tests	.18 (.09–.33)	1.0 (.80–1.0)	6.0* (2.9–12.4)	.84* (.70–1.1)

NOTE. Values are accuracy statistics (95% CIs). The best prediction rule based on the negative LR value is the presence of at least 2 positive predictor variables (prone instability test, aberrant movement, hypermobility, FABQ physical activity subscale >8).

*LR values estimated by adding 0.5 to each cell to avoid division by zero error.

fear, thus improving their disability level. Further work will be needed to clarify this issue. Finally, clinicians have suggested that hypermobility in the spine can be detected through mobility testing^{70,71} and is an indicator for stabilization exercises.⁷² The presence of hypomobility in the lumbar spine during mobility testing has been associated with the need for manipulation.⁴⁰ Mobility testing of the lumbar spine may be a useful clinical tool for determining the type of treatment most likely to benefit a particular patient.

LRs allow clinicians to determine the magnitude of the probability shift expected based on the result of an individual diagnostic test or a clinical prediction rule. The positive LR represents the change in odds favoring the condition of interest (stabilization success) when the criteria for the clinical prediction rule are met.⁵⁸ Therefore, a large positive LR is indicative of a larger shift in probability toward a positive response to stabilization. In this study, the prevalence, or pretest probability, of stabilization success was 33%. The posttest probability of stabilization success for a patient with at least 3 of the 4 predictor variables in the clinical prediction rule for success (positive LR, 4.0) would be 67%, suggesting that the patient may likely benefit from this stabilization program.

When examining the prediction rule for failure with the stabilization program, we chose to examine the negative LR because this statistic represents the change in odds favoring the condition of interest (some improvement with stabilization) when the criteria of the prediction rule are not met. In our study, the prevalence of some success with stabilization was 72%. The posttest probability of some success with stabilization for a patient with fewer than 2 of the 4 predictor variables in the clinical prediction rule (negative LR, .18) would be reduced to 32%, indicating that this person may be more likely to benefit from an alternative treatment approach. Conversely, the positive LR for this criterion is 6.3 and increases the likelihood of experiencing at least some improvement from 72% to 94%. When a cutoff score of 3 or more positive findings is used, the positive LR increases to 18.8, thereby increasing the probability of experiencing some improvement from 72% to 97% when 3 or more variables are present.

The stabilization program used in our study was standardized and each physical therapist involved followed the same protocol, but we acknowledge that clinician-based modifications to this program may produce superior results in certain situations. Although numerous types of stabilization exercises have been proposed, we believe the exercises used in this program are fairly representative and supported by the literature. We also attempted to control for any other types of biases that could temper the results of our study. First, the reference standard was applied to all subjects equally, and the examiner who judged the reference standard (patient response to stabilization) was blinded to the results of all diagnostic tests as well as to the clinical presentation of the subjects to avoid review bias. Additionally, the study design made it impossible for the clinician who judged the diagnostic test to be influenced by the result of the reference standard because it was not known until the 8-week time point. Finally, the inclusion of consecutive subjects with LBP from clinical practice should decrease the impact of spectrum bias and make the results generalizable to patients with LBP seeking outpatient PT treatment.

CONCLUSIONS

These results represent a preliminary step in the development of a clinical prediction rule for the use of stabilization exercises. We were able to identify several variables that appear to be important in determining which patients are likely or unlikely to benefit from a stabilization treatment approach. From the history and patient self-report, the variables of age, fear of physical activity, pain intensity, and frequency of previous episodes had some relation to treatment outcome. From the physical examination, the key variables appear to be SLR motion, lumbar mobility testing, aberrant motions during lumbar ROM, and the prone instability test. From these variables, we developed preliminary multivariate prediction rules. The results of this study need to be replicated in a separate sample for confirmation and ultimately examined in a randomized controlled trial before they can be recommended for widespread use.

APPENDIX 1: OPERATIONAL DEFINITIONS FOR PHYSICAL MEASURES

Physical Measures	Procedure
Range of motion	
Flexion ROM ³²	The patient stands and the inclinometer is held at T12-L1. The patient is asked to reach down as far as possible toward the toes while keeping the knees straight.
Extension ROM ³²	The patient stands and the inclinometer is held at T12-L1. The patient is asked to arch backward as far as possible.
Right and left side-bending ROM ³²	The patient stands with the inclinometer aligned vertically in line with the spinous processes of T9 and T12. The patient is asked to lean over to 1 side as far as possible with the fingertips reaching down the side of the thigh.

Physical Measures	Procedure
Right and left SLR ROM ³²	The patient is supine. The inclinometer is positioned on the tibial crest just below the tibial tubercle. The leg is raised passively by the examiner, whose other hand maintains the knee in extension. The leg is raised slowly to the maximum tolerated straight leg raise (not the onset of pain).
Muscle performance tests	
Side support test ⁴⁴	The patient is side lying with legs extended and the top foot in front of the lower foot. While resting on the lower elbow for support, the patient lifts the hips off the table with only the elbow and feet remaining in contact with the table. The patient is instructed to hold this position as long as possible. The test is done for both sides, and the performance time is recorded in seconds.
Extensor endurance test ⁴⁴	The patient is asked to lie prone while holding the sternum off the floor for as long as possible. A small pillow is placed under the lower abdomen to decrease the lumbar lordosis. The patient also needs to maintain maximum flexion of cervical spine and pelvic stabilization through gluteal contraction. The patient is asked to hold this position as long as possible not to exceed 5 minutes. The performance time is recorded in seconds.
Active sit-up test ³²	The patient is supine and is asked to flex the knees to 90° and place the soles of the feet flat on the surface. The examiner holds both feet down with 1 hand. The patient is instructed to reach up with the fingertips of both hands to touch (not hold) both knees and hold the position for 5 seconds. If the patient cannot maintain this position for 5 seconds, the test is positive.
Active bilateral SLR test ³²	The patient is supine and is asked to lift both legs together 6 inches (15.24 cm) off the examining surface and hold that position for 5 seconds. Both heels and calves should be cleared from the examining surface. If the patient cannot maintain this position for 5 seconds, the test is positive.
Special tests	
Posterior shear test ⁷	The patient is standing with arms across the lower abdomen. The examiner stands at 1 side of the patient and places 1 arm around the patient's abdomen, over the patient's crossed hands. The heel of the opposite hand is placed on the patient's pelvis for stabilization. The examiner produces a posterior force through the patient's abdomen and an anteriorly directed stabilizing force with the opposite hand. The test is repeated at all lumbar levels. A positive test is determined by the provocation of symptoms.
Prone instability test ⁴³	The patient lies prone with the body on the examining table and legs over the edge and feet resting on the floor. While the patient rests in this position, the examiner applies posterior to anterior pressure to the lumbar spine. Any provocation of pain is reported. Then the patient lifts the legs off the floor (the patient may hold table to maintain position) and posterior compression is applied again to the lumbar spine. If pain is present in the resting position but subsides in the second position, the test is positive.
Lumbar segmental testing for mobility ³⁷	The patient is prone. The L1 spinous process is contacted with the examiner's thenar eminence, and an anteriorly directed force is applied. The procedure is repeated at each lumbar level. Mobility is judged as hypermobile or not hypermobile.

References

- Waddell G. A new clinical model for the treatment of low-back pain. *Spine* 1987;12:632-44.
- Engel GL. The need for a new medical model: a challenge for biomedicine. *Science* 1977;196:129-36.
- Abenham L, Rossignol M, Gobeille D, Bonvalot Y, Fines P, Scott S. The prognostic consequences in the making of the initial medical diagnosis of work-related back injuries. *Spine* 1995;20:791-5.
- McKenzie RA. The lumbar spine: mechanical diagnosis and therapy. Waikanae (N Z): Spinal Publications Ltd; 1989.
- Spitzer WO. Approach to the problem. *Spine* 1987;12(Suppl):9-11.
- Von Korff M. Studying the natural history of back pain. *Spine* 1994;19(Suppl):S2041-6.
- 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;75:470-89.
- Bouter LM, van Tulder MW, Koes BW. Methodologic issues in low back pain research in primary care. *Spine* 1998;23:2014-20.
- Farfan HF, Gracovetsky S. The nature of instability. *Spine* 1984;9:714-9.
- Frymoyer JW, Selby DK. Segmental instability. Rationale for treatment. *Spine* 1985;10:280-6.
- Panjabi MM, Lydon C, Vasavada A, Grob D, Crisco JJ, Dvorak J. On the understanding of clinical instability. *Spine* 1994;19:2642-50.
- Knutsson F. The instability associated with disk degeneration in the lumbar spine. *Acta Radiol* 1944;25:593-609.
- Posner I, White AA 3rd, Edwards WT, Hayes WC. A biomechanical analysis of the clinical stability of the lumbar and lumbosacral spine. *Spine* 1982;7:374-89.
- White AA, Panjabi MM. Clinical biomechanics of the spine. 2nd ed. Philadelphia: JB Lippincott; 1990. p 23-45.
- Dupuis PR, Yong-Hing K, Cassidy JD, Kirkaldy-Willis WH. Radiographic diagnosis of degenerative lumbar spinal instability. *Spine* 1985;10:262-76.
- Boden SD, Wiesel SW. Lumbosacral segmental motion in normal individuals. Have we been measuring instability properly? *Spine* 1991;15:571-5.

17. Hayes MA, Howard TC, Gruel CR, Kopta JA. Roentgenographic evaluation of the lumbar spine flexion-extension in asymptomatic individuals. *Spine* 1989;14:327-31.
18. Fritz JM, Erhard RE, Hagen BF. Segmental instability of the lumbar spine. *Phys Ther* 1998;78:889-96.
19. Kirkaldy-Willis WH, Farfan HF. Instability of the lumbar spine. *Clin Orthop* 1982;May(165):110-23.
20. Paris SV. Physical signs of instability. *Spine* 1985;10:277-9.
21. Richardson C, Jull G, Hodges P, Hides J. Therapeutic exercise for spinal segmental stabilization in low back pain. Edinburgh: Churchill Livingstone; 1999.
22. Cholewicki J, McGill SM. Mechanical stability of the in vivo lumbar spine: implications for injury and chronic low back pain. *Clin Biomech (Bristol, Avon)* 1996;11:1-15.
23. Gardner-Morse MG, Stokes IA, Laible JP. Role of muscles in lumbar spine stability in maximum extensor efforts. *J Orthop Res* 1995;13:802-8.
24. Hides JA, Jull GA, Richardson CA. Long-term effects of specific stabilizing exercises for first-episode low back pain. *Spine* 2001;26:E243-8.
25. Hides JA, Richardson CA, Jull GA. Multifidus muscle recovery is not automatic after resolution of acute, first-episode low back pain. *Spine* 1996;21:2763-9.
26. O'Sullivan PB, Phytly GD, Twomey LT, Allison GT. Evaluation of specific stabilizing exercises in the treatment of chronic low back pain with radiologic diagnosis of spondylosis or spondylolisthesis. *Spine* 1997;22:2959-67.
27. Jensen MP, Turner JA, Romano JM. What is the maximum number of levels needed in pain intensity measurement? *Pain* 1994;58:387-92.
28. Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain* 1993;52:157-68.
29. Fritz JM, Irrgang JJ. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Phys Ther* 2001;81:776-88.
30. Davidson M, Keating JL. A comparison of five low back disability questionnaires: reliability and responsiveness. *Phys Ther* 2002;82:8-24.
31. Suarez-Almazor ME, Kendall C, Johnson JA, Skeith K, Vincent D. Use of health status measures in patients with low back pain in clinical settings. Comparison of specific, generic and preference-based instruments. *Rheumatology* 2000;39:783-90.
32. Waddell G, Somerville D, Henderson I, Newton M. Objective clinical evaluation of physical impairment in chronic low back pain. *Spine* 1992;17:617-28.
33. Nachemson A. Lumbar spine instability: a critical update and symposium summary. *Spine* 1985;10:290-1.
34. Ogon M, Bender BR, Hooper DM, et al. A dynamic approach to spinal instability. Part I. Sensitization of intersegmental motion profiles to motion direction and load condition by instability. *Spine* 1997;22:2841-58.
35. Cyriax J. Textbook of orthopaedic medicine. 6th ed. Baltimore: Williams & Wilkins; 1976. p 389.
36. 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.
37. Maher CG, Simmonds M, Adams R. Therapists' conceptualization and characterization of the clinical concept of spinal stiffness. *Phys Ther* 1998;78:289-300.
38. Binkley J, Stratford P, Gill C. Interrater reliability of lumbar accessory motion mobility testing. *Phys Ther* 1995;75:786-95.
39. Maher C, Adams R. Reliability of pain and stiffness assessments in clinical manual lumbar spine examination. *Phys Ther* 1994;74:801-11.
40. Flynn T, Fritz J, Whitman J, et al. A clinical prediction rule for classifying patients with low back pain who demonstrate short term improvement with spinal manipulation. *Spine* 2002;27:2835-43.
41. Beighton P, Horan F. Orthopedic aspects of Ehlers-Danlos syndrome. *J Bone Joint Surg Am* 1969;51:444-53.
42. Magee DJ. Orthopaedic physical assessment. 3rd ed. Philadelphia: WB Saunders; 1997. p 399.
43. McGill S, Childs A, Liebenson C. Endurance times for low back stabilization exercises: clinical targets for testing and training from a normal database. *Arch Phys Med Rehabil* 1999;80:941-4.
44. McGill SM. Low back exercises: evidence for improving exercise regimens. *Phys Ther* 1998;78:754-64.
45. McGill SM. Low back stability: from formal description to issues for performance and rehabilitation. *Exerc Sports Sci Rev* 2001;29:26-31.
46. Richardson CA, Jull GA. Muscle control-pain control: what exercises would you prescribe? *Man Ther* 1995;1:2-10.
47. Hodges PW, Richardson CA. Contraction of the abdominal muscles associated with movement of the lower limb. *Phys Ther* 1997;77:132-42.
48. Lee JH, Hoshino Y, Nakamura K, Kariya Y, Saita K, Ito K. Trunk muscle weakness as a risk factor for low back pain a 5-year prospective study. *Spine* 1999;24:54-7.
49. McGill SM. Estimation of force and extensor moment contributions of the disc and ligaments at L4-L5. *Spine* 1988;13:1395-402.
50. McGill S, Juker D, Kropf P. Quantitative intramuscular myoelectric activity of quadratus lumborum during a wide variety of lift tasks. *Clin Biomech (Bristol, Avon)* 1996;11:170-2.
51. Delitto A, Cibulka MT, Erhard RE, Bowling RW, Tenhula JA. Evidence for use of an extension-mobilization category in acute low back syndrome: a prescriptive validation pilot study. *Phys Ther* 1993;73:216-28.
52. Erhard RE, Delitto A, Cibulka MT. Relative effectiveness of an extension program and a combined program of manipulation and flexion and extension exercises in patients with acute low back syndrome. *Phys Ther* 1994;74:1093-100.
53. Fritz JM, George S. The use of a classification approach to identify subgroups of patients with acute low back pain: inter-rater reliability and short-term treatment outcomes. *Spine* 2000;25:114-21.
54. Nachemson AL. Advances in low-back pain. *Clin Orthop* 1985;Nov(200):266-78.
55. Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain* 1996;65:71-6.
56. Hagen MD. Test characteristics: how good is that test? *Med Decis Making* 1995;22:213-33.
57. Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. *J Chronic Dis* 1986;11:897-906.
58. Sackett DL. A primer on the precision and accuracy of the clinical examination. *JAMA* 1992;267:2638-44.
59. Bendebba M, Torgerson WS, Long DM. A validated, practical classification procedure for many persistent low back pain patients. *Pain* 2000;87:89-97.
60. Colle F, Rannou F, Revel M, Fermanian J, Poiraudou S. Impact of quality scales on levels of evidence inferred from a systematic review of exercise therapy and low back pain. *Arch Phys Med Rehabil* 2002;83:1745-52.
61. Ogon M, Bender BR, Hooper DM, et al. A dynamic approach to spinal instability. Part II. Hesitation and giving-way during inter-spinal motion. *Spine* 1997;22:2859-66.
62. van den Hoogen HM, Koes BW, van Eijk JT, Bouter LM. On the accuracy of history, physical examination, and erythrocyte sedimentation rate in diagnosing low back pain in general practice. *Spine* 1995;20:318-27.

63. Vroomen PC, de Krom M, Knottnerus JA. Predicting the outcome of sciatica at short-term follow-up. *Br J Gen Pract* 2002;52:119-23.
64. Vucetic N, Astrand P, Guntner P, Svensson O. Diagnosis and prognosis in lumbar disc herniation. *Clin Orthop* 1998;Apr(361):116-22.
65. Mannion AF, Kaser L, Weber E, Ryner A, Dvorak J, Muntener M. Influence of age and duration of symptoms on fibre type distribution and size of the back muscles in chronic low back pain patients. *Eur Spine J* 2000;9:273-81.
66. Savage RA, Millership R, Whitehouse CH, Edwards RH. Lumbar muscularity and its relationship with age, occupation and low back pain. *Eur J Appl Physiol* 1991;63:265-8.
67. Thompson LV. Effects of age and training on skeletal muscle physiology and performance. *Phys Ther* 1994;74:71-7.
68. Fritz JM, George SZ, Delitto A. The role of fear avoidance beliefs in acute low back pain: relationships with current and future disability and work status. *Pain* 2001;94:7-15.
69. Klenerman L, Slade PD, Stanley IM, et al. The prediction of chronicity in patients with an acute attack of low back pain in a general practice setting. *Spine* 1995;20:478-84.
70. Gronbald M, Hurri H, Kouri JP. Relationships between spinal mobility, physical performance tests, pain intensity and disability assessments in chronic low back pain patients. *Scand J Rehabil Med* 1997;29:17-24.
71. Maher CG, Latimer J, Adams R. An investigation of the reliability and validity of posteroanterior spinal stiffness judgments made using a reference-based protocol. *Phys Ther* 1998;78:829-37.
72. Maitland GD. *Vertebral manipulation*. 5th ed. Oxford: Butterworth Heinemann; 1986. p 74-6.