

# Identifying Subgroups of Patients With Acute/Subacute “Nonspecific” Low Back Pain

## Results of a Randomized Clinical Trial

Gerard P. Brennan, PhD, PT,\* Julie M. Fritz, PhD, PT, ATC,\* Stephen J. Hunter, MS, PT, OCS,\* Anne Thackeray, PT,\* Anthony Delitto, PhD, PT, FAPTA,† and Richard E. Erhard, DC, PT†

**Study Design.** Randomized clinical trial.

**Objective.** Compare outcomes of patients with low back pain receiving treatments matched or unmatched to their subgrouping based on initial clinical presentation.

**Summary of Background Data.** Patients with “nonspecific” low back pain are often viewed as a homogeneous group, equally likely to respond to any particular intervention. Others have proposed methods for subgrouping patients as a means for determining the treatment most likely to benefit patients with particular characteristics.

**Methods.** Patients with low back pain of less than 90 days’ duration referred to physical therapy were examined before treatment and classified into one of three subgroups based on the type of treatment believed most likely to benefit the patient (manipulation, stabilization exercise, or specific exercise). Patients were randomly assigned to receive manipulation, stabilization exercises, or specific exercise treatment during a 4-week treatment period. Disability was assessed in the short-term (4 weeks) and long-term (1 year) using the Oswestry. Comparisons were made between patients receiving treatment matched to their subgroup, *versus* those receiving unmatched treatment.

**Results.** A total of 123 patients participated (mean age, 37.7 ± 10.7 years; 45% female). Patients receiving matched treatments experienced greater short- and long-term reductions in disability than those receiving unmatched treatments. After 4 weeks, the difference favoring the matched treatment group was 6.6 Oswestry points (95% CI, 0.70–12.5), and at long-term follow-up the difference was 8.3 points (95% CI, 2.5–14.1). Compliance-only analysis of long-term outcomes yielded a similar result.

**Conclusions.** Nonspecific low back pain should not be viewed as a homogenous condition. Outcomes can be improved when subgrouping is used to guide treatment decision-making.

**Key words:** low back pain, acute/subacute, physical therapy, subgrouping, classification, randomized trial.  
**Spine 2006;31:623–631**

Low back pain (LBP) is a prevalent and costly condition, particularly in primary care.<sup>1,2</sup> Despite concentrated research efforts, few interventions have been identified as effective for patients with acute LBP.<sup>3,4</sup> One explanation offered for the failure to identify effective treatments is the lack of methods for subgrouping, or classifying, patients with “nonspecific” LBP in a manner that would help direct treatment decision-making.<sup>5</sup> Most clinicians managing patients with LBP perceive that subgroups exist, and are recognizable,<sup>6</sup> yet there is presently little unanimity in classification methods. The utility of a patho-anatomic classification method for the nonsurgical management of LBP is limited by an inability to identify a pathologic mechanism for most patients.<sup>7</sup> Emphasis in the development of classification methods for conservative care has therefore been placed on recognition of patterns of signs and symptoms from the clinical examination.<sup>6,8,9</sup>

The value of a classification method is based on its ability to improve clinical outcomes. Emerging evidence supports the hypothesis that classifying patients with LBP into subgroups based on signs and symptoms, and basing treatment on the subgrouping, produces better outcomes when compared with treatments not based on classification methods.<sup>10–13</sup> Fritz *et al*<sup>12</sup> compared the outcomes of patients with acute, work-related LBP randomized to classification treatment based on patterns of signs and symptoms, or treatment based on a clinical practice guideline with no attempt at subgrouping.<sup>14</sup> Patients randomized to classification treatment were placed in one of four subgroups, each with its own treatment approach (manipulation, stabilization, specific exercise, or traction). Patients receiving classification-based treatment had significantly better outcomes in terms of cost, disability, and return to work after 4 weeks and 1 year.<sup>12</sup> Childs *et al*<sup>13</sup> examined the validity of the manipulation subgroup. Patients with LBP were classified as either likely or not likely to respond to manipulation based on a pattern of signs and symptoms and then randomized patients to either a manipulation or exercise intervention. The treatment effect at 4 weeks and 6 months was greatest for the subgroup of patients classified as responders who received manipulation, demonstrating the validity of this subgroup.<sup>13</sup>

The study by Fritz *et al*<sup>12</sup> provides evidence that a classification-based management strategy generally produces better clinical outcomes than nonclassification management; however, this study used a pragmatic de-

From the \*Rehab Agency Intermountain Health Care, Salt Lake City, UT; and the †Department of Physical Therapy, University of Pittsburgh, Pittsburgh, PA.

Acknowledgment date: December 8, 2004. First revision date: February 8, 2005. Second revision date: March 23, 2005. Acceptance date: April 4, 2005.

Supported by a research grant from the Deseret Foundation.

The manuscript submitted does not contain information about medical device(s)/drug(s).

Foundation funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

Address correspondence and reprint requests to Julie M. Fritz, PhD, PT, ATC, 520 Wakara Way, Salt Lake City, UT 84108; E-mail: julie.fritz@hsc.utah.edu

sign that randomized patients to either guideline-based treatment (low-stress aerobic exercise and advice to remain active) or classification-based treatment (manipulation, specific exercise, stabilization exercise, or traction). Given this design, it is possible that improved outcomes were the result of a better set of treatments, and not the classification process *per se*. Only one prior study has looked specifically at the interaction between the treatment received and one classification subgroup (manipulation).<sup>13</sup> The purpose of this study was to determine if patients with LBP would demonstrate greater and more rapid functional improvement based on the initial treatment received (regardless of classification), the classification subgrouping (regardless of treatment), or the interaction of the two factors.

## ■ Methods

**Subjects.** Patients referred to physical therapy with a chief complaint of LBP were considered for inclusion. Primary recruitment occurred at one clinic between January 1, 2000 and July 1, 2003. Additional recruitment occurred at two other clinics between January 1, 2002 and September 1, 2002. Each clinic was located in Utah and affiliated with Intermountain Health Care System. The study was approved by the Institutional Review Board of Intermountain Health Care, and all patients provided informed consent before participation.

Patients between 18 and 65 years with a primary complaint of LBP of less than 90 days, with or without referral into the lower extremity, and an Oswestry disability score  $\geq 25\%$  were eligible. Exclusion criteria were a visible lateral shift or acute kyphotic deformity, signs of nerve root compression (positive straight leg raise test *and* reflex or strength deficits), any red flags indicating a serious pathology such as spinal neoplasm, infection, or fracture, an inability to reproduce any symptoms with lumbar spine active range of motion (AROM) or palpation, current pregnancy, or prior surgery to the lumbar and/or sacral region.

**Baseline Examination.** Before randomization, patients completed a set of questionnaires and underwent a standardized examination. Baseline and follow-up examinations were conducted by a physical therapist who remained blind to the treatment group assignment. Demographic information including age, sex, prior history of LBP, and duration and location of current symptoms were collected. An 11-point rating scale (0 = no pain to 10 = worst imaginable pain) was used to assess current pain intensity.<sup>15</sup> Fear-avoidance beliefs were assessed through the Fear Avoidance Beliefs Questionnaire (FABQ).<sup>16</sup> The two subscales of the FABQ contain 11 items, each scored 0 to 6, with higher numbers indicating increased levels of fear-avoidance beliefs. Both the 7-item work subscale (FABQW) and the 4-item physical activity subscale (FABQPA) were assessed. Previous studies have found strong relationships between fear-avoidance beliefs and disability and work loss in patients with acute and chronic LBP.<sup>17-19</sup> The Modified Oswestry Questionnaire (OSW) was used to assess disability related to LBP. This modified version replaces the sex life item with an employment/homemaking item and has been found to be reliable, valid, and sensitive to change.<sup>20</sup>

The physical examination included questions regarding aggravating and relieving activities and prior history of LBP.

Tests and measurements included lumbar AROM,<sup>21</sup> judgment of centralization or peripheralization with AROM including the movements of flexion, extension, side-bending, pelvic translations, and repeated extension performed with the patient standing, as well as sustained prone extension, and repeated flexion with the patient seated.<sup>22</sup> We have found excellent interrater reliability for judgments of centralization/peripheralization among therapists of varying levels of clinical experience.<sup>23</sup> Aberrant movements occurring during AROM, including instability catch, painful arc of movement, or tortuous return from the flexed position, were recorded, and the segmental instability test, a special test thought to indicate a need for stabilizing exercises, was performed.<sup>24</sup> The mobility of each level of the lumbar spine was assessed using posterior-anterior forces applied with the patient prone.<sup>25</sup> Each segment was judged to be normal, hypomobile, or hypermobile. Further explanation of the operational definitions and reliability of these procedures is published elsewhere.<sup>26</sup>

**Treatment.** A random number generator was used to generate a randomization list before initiation of the study. The list was maintained by the secretarial staff of the participating clinics. Following the baseline examination, the patient was referred to a different therapist to begin treatment. Before the first treatment session, the secretarial staff consulted the randomization list and assigned the patient to one of three groups: manipulation, specific exercise, or stabilization. This assignment determined the treatment the patient would receive during the initial stage of therapy. Once sufficient progress was made, patients were progressed to a general exercise program (Stage II). Methods such as heat, ice, and electrical stimulation could be used in any group at the discretion of the treating therapist. All patients were scheduled for treatment twice weekly for 4 weeks for a maximum of eight sessions.

**Manipulation Treatment Group.** Patients randomized to the manipulation group were treated with manual therapy techniques that could include thrust manipulation, or low amplitude mobilization procedures directed to the lumbosacral region, along with instruction in a lumbar AROM exercise. The therapist performing the treatment was permitted to reexamine the patient and could choose one of two manual therapy techniques. The choice of which technique to use was left to the therapists' discretion, but one of the two techniques had to be used. In the first technique, the patient was supine, with the lumbar spine placed into side-bending and rotation to the opposite direction. The therapist delivered a force through the patient's pelvis in a posterior and inferior direction. For the second technique, the patient was side-lying. The lumbar spine was positioned in either flexion or extension followed by rotation in an attempt to isolate forces to a particular spinal level. The therapist delivered the force through the patient's pelvis and trunk. The choice of technique was left to the discretion of the therapist. The AROM exercise was performed by instructing the patient to alternately flex and extend the lumbar spine while in a quadruped position.

**Specific Exercise Treatment Group.** Patients in the specific exercise group received instruction in repeated ROM exercises into either lumbar flexion or extension. All patients in this group had to be treated with directional exercises; however, the direction of the exercise was determined by the treating therapist based on a reassessment of the patient's response to movement testing and symptom response to positions of sitting,

standing, or walking. Flexion exercises were used for patients who centralized with or had a preference for flexion movements or positions (*i.e.*, sitting), whereas extension exercises were used for patients who centralized or had a preference for extension (*i.e.*, standing or walking). Either flexion or extension exercises were used, but not both. Flexion exercises were performed with the patient sitting, supine, or quadruped. Extension exercises were performed in prone, using prone on elbows or prone press-up activities.

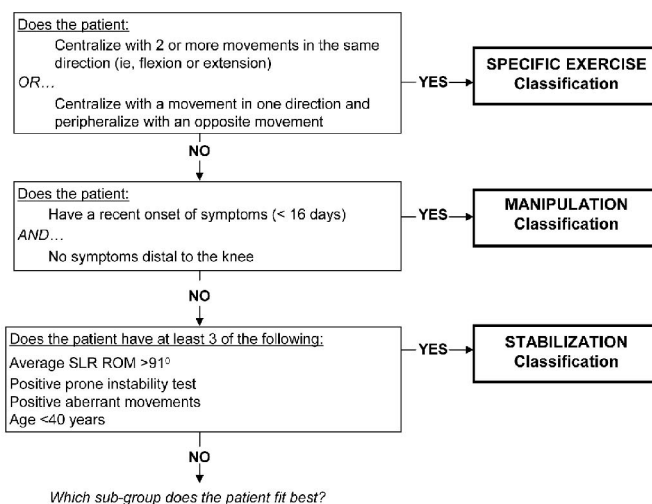
**Stabilization Treatment Group.** Patients in the stabilization group were treated with a program of trunk strengthening and stabilization exercises. Patients were instructed to perform abdominal bracing exercises in supine and quadruped positions, progressing to more functional positions and activities as described by Richardson and Jull.<sup>27</sup> Patients were also instructed in alternating arm and leg extension exercises in quadruped to strengthen the lumbar extensor muscles.<sup>28</sup> Strengthening for the oblique abdominals included curl-up and side support exercises.<sup>29</sup>

**Treatment Progression.** This study used a pragmatic approach to treatment progression. The treating therapists could select only those treatments permitted based on the patient’s treatment group, but the therapist used clinical judgment to determine exercise dosage for individual patients. Treatment was based on the randomized group assignment until the patient progressed into a second (subacute) stage. If the patient did not progress into Stage II, treatment continued to be based on the initial group assignment throughout the 4-week period. The criteria to progress to Stage II were based on changes in the OSW score, which was collected at the beginning of each treatment session. If the OSW score was reduced to <20%, or a reduction of 33% from the score recorded at the first treatment session was achieved, the patient was progressed to Stage II treatment. Progression of treatment was permitted to more accurately reflect clinical practice, in which treatment is typically altered and progressed as improvements are made.<sup>30</sup>

Once in treatment Stage II, all patients received a general exercise program in keeping with evidence-based recommendations advocating an active, multimodal exercise approach for patients who have progressed beyond the acute stage of LBP.<sup>31,32</sup> The Stage II program consisted of a low-stress aerobic exercise program using a treadmill or stationary bike and exercises to address any muscle strength or flexibility impairments identified by the treating therapist (*e.g.*, hamstring stretches, quadriceps strengthening, *etc.*). Based on evidence suggesting a reduction in recurrence rates with strength and stabilization exercises,<sup>33</sup> patients not originally in the stabilization group were instructed to begin these exercises. Patients could also continue any Stage I exercises they were performing based on their randomized treatment group.

**Follow-up Examinations.** Two follow-up examinations were conducted at which the OSW was readministered. The short-term follow-up occurred at the completion of treatment, approximately 4 weeks after the baseline assessment. The long-term follow-up was conducted by mail approximately 1 year following the completion of treatment.

**Patient Subgrouping.** After completion of study, the baseline examination data from each patient were used to classify each patient into one of three classification subgroups. The classification was made based on previous work describing a decision-making scheme that seeks to place patients with LBP into one of four subgroups (manipulation, stabilization, specific exercise, and traction) by considering the signs and symptoms from the clinical examination (Figure 1).<sup>8,26,34</sup> The traction subgroup was not considered in this study because patients with signs of nerve root compression were excluded. The baseline signs and symptoms of each patient were reviewed by two physical therapists with extensive clinical experience using this decision-making scheme. If the two experts agreed on the subgrouping decision, the patient was considered to fit that subgroup. If the two experts did not agree, a third expert made a



MANIPULATION		STABILIZATION		SPECIFIC EXERCISE	
Factors favoring	Factors against	Factors favoring	Factors against	Factors favoring	Factors against
More recent onset of symptoms	Symptoms below the knee	Younger age	Discrepancy in SLR ROM (>10°)	Strong preference for sitting or walking	Low back pain only (no distal symptoms)
Segmental hypomobility	Increasing episode frequency	Positive prone instability test	Low FABQ scores (FABQPA < 9)	Centralization with motion testing	Status Quo with all movements
LBP only (no distal symptoms)	Peripheralization with motion testing	Aberrant motions present		Peripheralization in direction opposite centralization	
Low FABQ scores (FABQW <19)	No pain with mobility testing	Greater SLR ROM			
		Hypomobility with spring testing			
		Increasing episode frequency			
		3 or more prior episodes			

Figure 1. Decision-making algorithm used for subgrouping patients.

judgment and the patient was considered to fit the subgroup agreed on by two of three experts. The expert therapists were blind to the treatment group assignment and outcome. Further detail on the classification procedures is provided elsewhere.<sup>26</sup>

Once the classification subgroup was determined, the subgrouping was compared with the randomized treatment group assignment for each patient. If the treatment group to which the patient was randomized matched the classification subgroup, the patient was categorized as having received the matched treatment. If the randomized treatment group did not correspond to the classification subgroup, the patient was categorized as unmatched.

**Data Analysis.** Baseline variables were compared between groups using independent *t* tests, Mann-Whitney U tests, or  $\chi^2$  tests of independence based on the nature of the data. To examine the principle hypothesis of the study, a three-way repeated-measures analysis of variance was performed with treatment group (manipulation *vs.* specific exercise *vs.* stabilization) and classification subgroup (manipulation *vs.* specific exercise *vs.* stabilization) as between-subject variables and time (baseline, 4 weeks, and 1 year) as the within-subject variable. The dependent variable was disability (OSW). The hypothesis of interest was the three-way interaction, and the two-way interactions between time and treatment group, and between time and classification subgroup. We hypothesized that outcome over time would not differ based on the randomized treatment group, or the classification subgroup, but would depend on the interaction between the treatment group and classification subgroup, such that patients randomized to treatment matched to the subgroup would have better outcomes than patients randomized to an unmatched treatment. This hypothesis would be supported if the three-way interaction was significant, but the two-way interactions were not. Pairwise *post hoc* comparisons were performed at each follow-up period to further explore any significant interaction terms.

We further examined the impact of receiving matched treatment by comparing the proportions of patients progressing to Stage II treatment, and the number of treatment sessions re-

quired to reach Stage II based on the randomized treatment group, and based on matching status using separate  $\chi^2$  tests for the proportions reaching Stage II, and Kruskal-Wallis tests for the number of treatment sessions. We hypothesized that no differences would exist based on randomized treatment group or classification subgroup, but patients receiving matched treatments would progress to Stage II in higher proportions, and in fewer treatment sessions than patients receiving unmatched treatment. Analysis was performed using intention-to-treat principles, with the last available OSW score carried forward for any missing data. An alpha level of 0.05 was used for all comparisons.

**Sample Size Determination.** Sample size calculation was based on detecting the three-way interaction using the 4-week OSW score as the dependent variable with an alpha value of 0.05. A within-cell standard deviation of 14 points on the OSW and a correlation across repeated measures of the OSW of 0.30 were presumed from prior work.<sup>34</sup> Given these estimates, 120 patients would be sufficient to detect a minimum clinically important difference of 6 points on the OSW<sup>20</sup> (effect size = 0.43) with 80% power using a two-tailed hypothesis.

## ■ Results

A total of 1,052 patients with LBP were referred to participating clinics during the recruitment period, of which 268 were eligible for participation. The most common reasons for exclusion were the duration of symptoms ( $n = 237$ ) or age ( $n = 203$ ) (Figure 2). A total of 123 eligible patients provided consent and were enrolled; 40 were randomized to manipulation, 46 to stabilization, and 37 to the specific exercise treatment group.

The expert therapists agreed on the classification subgroup for 102 patients (83%). A third expert was used to determine the classification subgroup for the remaining 21 patients. The classification subgroup was manipulation for 59 patients (48%), specific exercise for 34 pa-

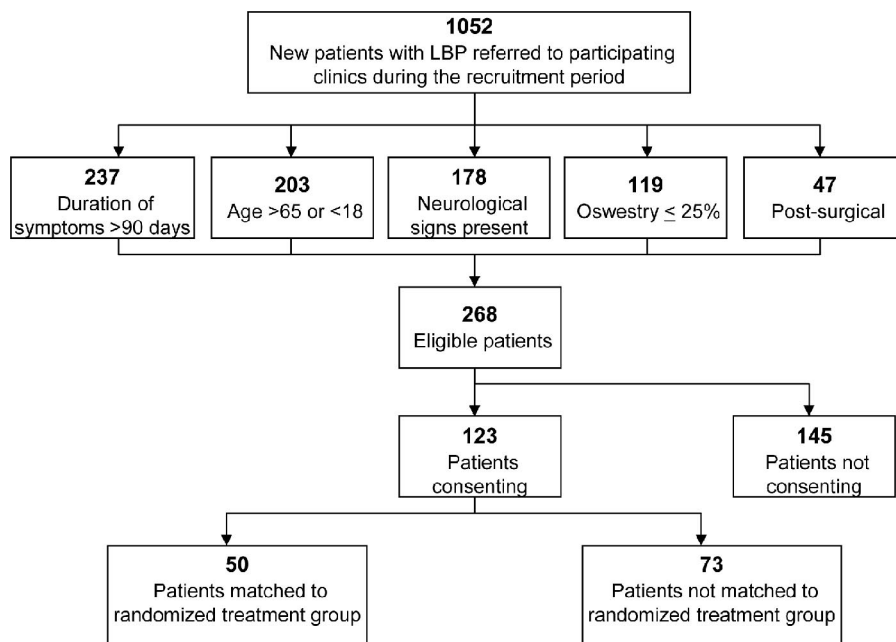


Figure 2. Patient recruitment flow diagram.

**Table 1. Baseline Characteristics of Study Participants**

Variable	All Patients (n = 123)	Patients Receiving Matched Treatment (n = 50)	Patients Not Receiving Matched Treatment (n = 73)	Significance (P)
Age (yr)	37.7 (10.7)	38.4 (10.0)	37.2 (11.1)	0.51
Sex (% female)	45	34	52	0.049
Education level (% college graduates)	24	24	24	0.87
Prior history of low back pain (%)	67	70	64	0.52
Symptoms distal to the knee (%)	22	28	18	0.16
Duration of current symptoms (median days, interquartile range)	16 (10, 41)	15 (10, 32)	17 (10, 54)	0.53
Missed work/school for this episode of low back pain (%)	55	52	56	0.21
FABQ: physical activity scale	16.4 (4.9) (n = 115)	16.5 (5.3) (n = 46)	16.3 (4.7) (n = 69)	0.87
FABQ: work scale	14.6 (10.5) (n = 112)	15.7 (10.7) (n = 45)	13.8 (10.4) (n = 67)	0.37
Oswestry disability score	43.3 (11.5)	44.9 (11.2)	42.2 (11.7)	0.20
Pain score	5.4 (2.0) (n = 116)	5.2 (2.1) (n = 46)	5.6 (1.9) (n = 70)	0.33

Values represent the mean (SD), except where noted otherwise.

tients (28%), and stabilization for 30 patients (24%). Comparison of the classification subgroup with the randomized treatment group resulted in 50 patients (41%) categorized as receiving matched treatment and 73 (59%) receiving unmatched treatment. Of the 50 patients receiving matched treatment, 22 were manipulation, 14 were specific exercise, and 14 were stabilization.

Patients receiving matched and unmatched treatments were equivalent for all baseline characteristics (Table 1) except sex, with a greater percentage of females in the unmatched group (52% vs. 34%,  $P = 0.049$ ). Ten patients (3 matched, 7 unmatched) did not return after the baseline examination. These patients were included in the analysis by carrying forward the baseline OSW score. Among patients attending treatment, there were no differences in the median number of sessions attended, or

days from baseline to 4-week reexamination (Table 2). The median number of sessions attended by the matched group was 6.5, and 80% attended at least four sessions. In the unmatched group, the median number of sessions was seven, and 77% attended at least four sessions. Eighty-one patients (66%) completed the long-term follow-up, with no differences in the median number of days between baseline and follow-up or the proportions of patients with completed follow-up between patients receiving matched or unmatched treatments (Table 2). There were no significant differences between those with complete or incomplete long-term follow-up with respect to age, sex, duration of symptoms, baseline pain, OSW, or FABQ scores ( $P > 0.05$ ). The proportion of matched versus unmatched patients did not differ between those with complete or incomplete long-term follow-up ( $P >$

**Table 2. Treatment Characteristics for Patients Receiving Matched or Unmatched Treatments**

Variable	Patients Receiving Matched Treatment (n = 50)	Patients Not Receiving Matched Treatment (n = 73)	Significance (P)
Median therapy sessions attended (interquartile range)	6.5 (4, 8)	7 (4, 8)	0.71
Time from baseline to 4-week examination (median, days)	32 (n = 47)	31 (n = 66)	0.62
Change in Oswestry from baseline to 4-week examination	29.9 (15.6)	23.3 (16.6)	0.03
Percent progressing to Stage II treatment	78%	60%	0.039
Treatment sessions needed to progress to Stage II (includes only patients progressing to Stage II)	2.2 (1.3) (n = 39)	2.8 (2.0) (n = 44)	0.070
Percent with completed long-term follow-up examination	60%	70%	0.26
Time from baseline to long-term examination (median, days)	421 (n = 30)	414 (n = 51)	0.77
Change in Oswestry from baseline to long-term examination (intention-to-treat analysis)	27.9 (16.0)	19.6 (16.0)	0.006
Change in Oswestry from baseline to long-term examination (compliers-only analysis)	33.3 (13.9) (n = 30)	26.1 (16.4) (n = 51)	0.049

Values represent the mean (SD), except where noted otherwise.

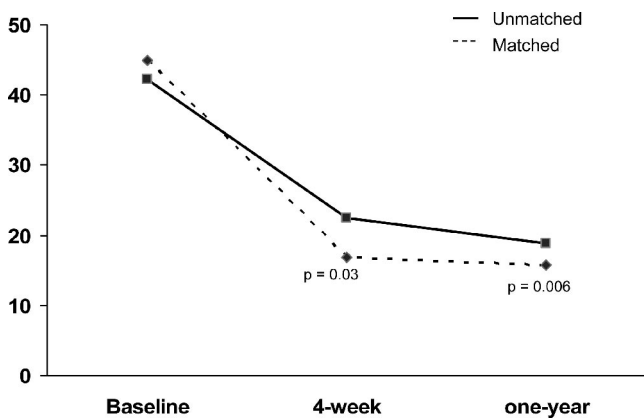


Figure 3. Oswestry scores for patients receiving matched or unmatched treatments (intention-to-treat analysis,  $P$  values represent differences between the baseline and follow-up scores).

0.05). Patients with complete long-term follow-up did have lower 4-week OSW scores (17.3 vs. 25.0,  $P = 0.02$ ). Patients with incomplete long-term follow-up were included in the analysis by carrying forward the 4-week OSW score.

The three-way interaction between randomized treatment group, classification subgroup, and time was significant ( $P = 0.016$ ); the two-way interactions between randomized treatment group and time ( $P = 0.26$ ), and classification subgroup and time ( $P = 0.13$ ) were not significant. Further exploration of the three-way interaction revealed that patients receiving matched treatments experienced greater change on the OSW than patients receiving unmatched treatments at both the 4-week (mean difference, 6.6 points; 95% confidence interval [CI], 0.70–12.5,  $P = 0.029$ ), and 1-year follow-up (mean difference, 8.3 points; 95% CI, 2.5–14.1,  $P = 0.006$ ) (Figure 3). Because of the large proportion of patients lost to the long-term follow-up, we conducted a compliers-only analysis, including only patients completing a 1-year OSW score. Among compliers, patients receiving matched treatments experienced greater long-term change on the OSW (mean difference, 7.1 points;

95% CI, 0.02–14.3,  $P = 0.049$ ) (Table 2). The OSW scores patients at each follow-up based on randomized treatment group, classification subgroup, and matching status are given in Table 3.

There were no differences in the proportions of patients progressing to Stage II based on randomized treatment group or classification subgroup (Table 4). Patients receiving matched treatments were more likely to progress to Stage II of treatment, but no difference was found in the average number of visits required to progress to Stage II (Table 2).

## Discussion

Effective management of LBP has remained largely an enigma, with large variations in practice and less than optimal outcomes.<sup>35,36</sup> Attempts to overcome ineffectiveness and inefficiency have centered on developing evidence-based practice guidelines, yet implementation of guidelines for LBP management has proven challenging and their effectiveness in improving outcomes is uncertain.<sup>37–40</sup> One explanation may be that primary studies, and resultant guidelines, have conceptualized LBP as a mostly homogeneous condition once red flags and neurologic compromise are excluded, while clinicians report further tailoring management to fit individual patient characteristics.<sup>41</sup> Clinicians and researchers alike recognize the irrationality of the expectation that patients with nonspecific LBP would all benefit from any one type of treatment, yet this expectation appears implicit in the design of many studies and practice guidelines. Many researchers have called for future research to incorporate subgrouping methods<sup>5,42,43</sup>; however, such methods have not been fully elucidated.

The subgrouping method used in this study was originally derived from opinions and observations of various clinical experts<sup>8,44</sup>; however, subsequent research has provided evidence for determining which signs and symptoms relate to success with manipulation,<sup>45,46</sup> stabilization,<sup>2,3,47</sup> and specific exercise interventions.<sup>48–50</sup> The results of this study support the hypothesis that clin-

**Table 3. Oswestry Disability Scores at Each Follow-Up Period Based on Randomized Treatment Group, Classification Subgroup, and the Interaction Between Treatment and Classification**

Grouping Factor	Initial Oswestry (SD)	Four-Week Oswestry (SD)	Long-Term Oswestry (SD)	Significance ( $P$ )*
Randomized treatment group				
Manipulation	44.0 (12.0)	17.9 (17.6)	16.8 (18.5)	0.37
Specific exercise	42.1 (10.7)	20.6 (16.4)	14.8 (14.8)	
Stabilization exercise	43.6 (11.9)	21.9 (17.0)	20.5 (18.1)	
Classification subgroup				
Manipulation	42.2 (10.9)	18.1 (15.9)	14.9 (15.1)	0.19
Specific exercise	47.0 (12.0)	21.2 (17.7)	22.2 (17.8)	
Stabilization exercise	41.2 (11.5)	23.4 (18.1)	17.6 (20.1)	
Interaction between treatment and classification				
Matched	44.9 (11.1)	16.9 (17.3)	15.7 (17.2)	0.013
Unmatched	42.2 (11.7)	22.5 (16.4)	18.8 (17.5)	

\*Significance represents the  $P$  value from repeated-measures ANOVA using the grouping variable (randomized group, classification subgroup, or matching status) as the between-subjects factor and time as the within-subjects factor.

**Table 4. Proportions of Patients Progressing to Stage II Treatment Based on Randomized Treatment Group, Classification Subgroup, and the Interaction Between Treatment and Classification**

Grouping Factor	No. (%) Progressing to Stage II	No. (%) Not Progressing to Stage II	Significance (P)*
Randomized treatment group			
Manipulation	30 (75)	10 (25)	0.35
Specific exercise	22 (60)	15 (40)	
Stabilization exercise	31 (67)	15 (33)	
Classification subgroup			
Manipulation	41 (69)	18 (31)	0.60
Specific exercise	24 (71)	10 (29)	
Stabilization exercise	18 (60)	12 (40)	
Interaction between treatment and classification			
Matched	39 (78)	11 (22)	0.039
Unmatched	44 (60)	29 (40)	

\*Significance represents the P value from the  $\chi^2$  test of significance.

ical outcomes can be improved when the initial treatment provided is matched to a patient’s signs and symptoms using this subgrouping method. These results, and those of previous studies,<sup>12,13,50</sup> indicate that while treatments such as manipulation and various forms of exercise may show small, marginally significant treatment effects when delivered to heterogeneous samples, these same treatments may demonstrate clinically important effects when applied to more homogeneous groups of patients with LBP.

In this study, the randomized treatment group of the patient determined the initial intervention, and once progress was made a more general program was used. This approach reflects clinical practice, which typically involves reassessment and modification of treatment as a patient’s signs and symptoms change. In addition, the association between signs and symptoms and function in patients with LBP weakens with increasing time from onset,<sup>51</sup> which may make subgrouping less relevant as patients’ status becomes more stable. Patients initially receiving matched interventions in this study experienced greater reductions in disability of a magnitude that was both statistically significant and clinically meaningful,<sup>20,52</sup> averaging approximately 20% greater reductions than patients receiving unmatched interventions. Although the completeness of the long-term follow-up precludes definitive conclusions, the persistence of improvement at long-term follow-up among patients initially receiving matched interventions supports the premise that using more effective initial treatments has enduring implications. Given the increasing likelihood of persistent disability among patients who fail to recover quickly,<sup>53</sup> the initial treatment phase may be the “window of opportunity” for clinicians. Improved decision-making through subgrouping during the initial treatment phase may therefore have important long-term consequences. Subgrouping may be less important among patients with more chronic symptoms, or may be more effectively performed on the basis of psychosocial factors; however, more research is needed in these areas.<sup>54</sup>

We used a pragmatic approach to the treatment delivered within each subgroup. The study design dictated the initial type of treatment that would be received, but the treating therapist was permitted to determine the dosage of treatment. This was done in an attempt to more accurately reflect clinical practice and permit the treating therapist to use judgment within the confines of the randomized treatment group, instead of using strict treatment protocols. It is possible that the treatment effect of the matched patients would have been greater had strict protocols been used; however, it is also possible that the effect would have been smaller if the treating therapists’ clinical judgment had been removed completely. The results of this study support the value of the decision-making scheme used to determine the type of initial treatment to use. Further research is needed to determine the most effective protocols within a particular type of treatment.

A higher percentage of the patients receiving unmatched treatment were female. There is presently no evidence of sex differences in response to the treatments used in this study; however, the sex difference may have impacted on the results.

All patients in this study received physical therapy. It is probable that a subgroup of individuals with LBP exists who do not require referral or any specific intervention other than encouragement to remain active.<sup>55</sup> The results of this study cannot address the characteristics defining this potential subgroup. Many patients with LBP seen in primary care, however, do experience persistent, even disabling, symptoms,<sup>56,57</sup> making research into treatment decision-making necessary. This study focused on patients with symptoms of less than 90 days duration, with moderate disability, who did not have signs of nerve root compression. For patients fitting this profile, evidence supports the effectiveness of decision-making that matches treatments to signs and symptoms. Patients with more mild disability or more chronic symptoms will require alternative decision-making schemes, and further research is needed in these areas. The exclusion of patients with signs of nerve root compression likely affected the distribution of classifi-

cation subgroups within this study. In our previous work that included patients with signs of nerve root compression, a higher proportion of patients classified as specific exercise was found.<sup>12</sup>

The pragmatic approach used in this study may have resulted in smaller treatment effects than other recent trials that examined the outcomes of patients receiving matched *versus* unmatched treatments. The treatment effect observed in this study after 4 weeks was 0.40. Long and Donelson<sup>50</sup> examined patients exhibiting a directional preference (similar to patients classified as specific exercise in this study) and found treatment effects of approximately 0.50 to 0.70 for the outcome of disability after 2 weeks of treatment for patients receiving the matched directional exercise *versus* a general exercise approach, or a directional exercise approach in the opposite direction. Childs *et al*<sup>13</sup> reported on the subgroup of patients fitting a manipulation classification with a treatment effect greater than 1.0 for the outcome of disability after 4 weeks of treatment comparing patients receiving the manipulation *versus* patients receiving stabilization exercise. Each of these studies points to the importance of matching treatments to the characteristics of the patient and the need for further research into optimal methods for determining subgroups and delivering the treatments.

## ■ Conclusion

The results of this study build on prior findings suggesting that “nonspecific LBP” is actually a heterogeneous condition. Meaningful subgroups of patients can be identified based on signs and symptoms from the clinical examination. Better clinical outcomes can be achieved when subgrouping is used to guide treatment decisions.

## Acknowledgments

The authors thank Eric Passey, Bradley Zollinger, Locke Ettinger, Virgil Beck, and Shannon Clifford for assisting with the conduct of this study.

## ■ Key Points

- Identification of a pathoanatomical cause is elusive for many patients with low back pain, invoking a diagnostic label of “nonspecific” low back pain.
- Subgrouping patients with “nonspecific” low back pain based on signs and symptoms has been recommended by researchers and clinicians and supported by recent evidence as a means to improve outcomes.
- This study placed patients with acute/subacute low back pain into one of three treatment subgroups based on their initial signs and symptoms (manipulation, stabilization or specific exercise), then randomized patients to receive one of the three treatments.

- The short- and long-term outcomes did not differ based on the randomized treatment group, or the subgroup, but depended on the interaction between treatment group and subgroup, such that patients receiving the treatment matched to their subgroup had better outcomes than patients randomized to an unmatched treatment.
- Developing methods to subgroup patients with “nonspecific” low back pain can improve the outcomes of care.

## References

1. Deyo RA, Phillips WR. Low back pain: a primary care challenge. *Spine* 1996;21:2826–32.
2. Luo X, Pietrobon R, Sun SX, et al. Estimates and patterns of direct health care expenditures among individuals with back pain in the United States. *Spine* 2004;29:79–86.
3. van Tulder MW, Malmivaara A, Esmail R, et al. Exercise therapy for low-back pain. *Cochrane Database of Systematic Reviews*. 2004;3.
4. Assendelft WJ, Morton SC, Yu EI, et al. Spinal manipulative therapy for low-back pain. *Cochrane Database of Systematic Reviews*. 2004;3.
5. Bouter LM, van Tulder MW, Koes BW. Methodologic issues in low back pain research in primary care. *Spine* 1998;23:2014–20.
6. Kent P, Keating JL. Do primary-care clinicians think that nonspecific low back pain is one condition? *Spine* 2004;29:1022–31.
7. Abenheim L, Rossignol M, Gobeille D, et al. The prognostic consequences in the making of the initial medical diagnosis of work-related back injuries. *Spine* 1995;20:791–5.
8. Delitto A, Erhard RE, Bowling RW. A treatment-based classification approach to low back syndrome: identifying and staging patients for conservative management. *Phys Ther* 1995;75:470–89.
9. Spitzer WO. Approach to the problem. In: Scientific approach to the assessment and management of activity-related spinal disorders: a monograph for clinicians. *Spine* 1987;12(suppl):9–11.
10. Delitto A, Cibulka MT, Erhard RE, et al. Evidence for use of an extension-mobilization category in acute low back syndrome: a prescriptive validation pilot study. *Phys Ther* 1993;73:216–28.
11. 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.
12. Fritz JM, Delitto A, Erhard RE. Comparison of a classification-based approach to physical therapy and therapy based on clinical practice guidelines for patients with acute low back pain: a randomized clinical trial. *Spine* 2003;28:1363–72.
13. Childs JD, Fritz JM, Flynn TW, et al. Validation of a clinical prediction rule to identify patients with low back pain likely to benefit from spinal manipulation. *Ann Intern Med* 2004;141:920–8.
14. Bigos S, Bowyer O, Braen G, et al. *Acute Low Back Problems in Adults* [AHCPR Publication No. 95-0642]. Rockville, MD: Agency for Health Care Policy and Research, Public Health Service, U.S. Department of Health and Human Services, 1994.
15. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. *Spine* 2005;30:686–96.
16. Waddell G, Newton M, Henderson I, et al. 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.
17. Woby SR, Watson PJ, Roach NK, et al. Are changes in fear avoidance beliefs, catastrophizing, and appraisals of control, predictive of changes in chronic pain and disability? *Eur J Pain* 2004;8:201–10.
18. 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.
19. Crombez G, Vlaeyen JW, Heuts PH, et al. Pain-related fear is more disabling than fear itself: evidence on the role of pain-related fear in chronic back pain disability. *Pain* 1999;80:329–39.
20. Fritz JM, Irrgang JJ. A comparison of a Modified Oswestry Disability Questionnaire and the Quebec Back Pain Disability Scale. *Phys Ther* 2001;81:776–88.

21. Waddell G, Somerville D, Henderson I, et al. Objective clinical evaluation of physical impairment in chronic low back pain. *Spine* 1992;17:617-28.
22. McKenzie RA. *The Lumbar Spine: Mechanical Diagnosis and Therapy*. Waikanae, New Zealand: Spinal Publications, 1989.
23. Fritz JM, Delitto A, Vignovic M, et al. Inter-rater reliability of judgments of the centralization phenomenon and status change during movement testing in patients with low back pain. *Arch Phys Med Rehabil* 2000;81:57-61.
24. Hicks GE, Fritz JM, Delitto A, et al. The reliability of clinical examination measures used for patients with suspected lumbar segmental instability. *Arch Phys Med Rehabil* 2003;84:1858-64.
25. Maitland GD. *Vertebral Manipulation*, 5th ed Oxford, UK: Butterworth Heinemann, 1986.
26. Fritz JM, Brennan GP, Clifford SN, et al. An examination of the reliability of a classification algorithm for sub-grouping patients with low back pain. *Spine* 2006;31:77-82.
27. Richardson CA, Jull GA. Muscle control-pain control: what exercises would you prescribe? *Man Ther* 1995;1:2-10.
28. McGill SM. Low back exercises: evidence for improving exercise regimens. *Phys Ther* 1998;78:754-64.
29. Jucker D, McGill S, Kropf P, et al. Quantitative intramuscular myoelectric activity of lumbar portions of psoas and the abdominal wall during a wide variety of tasks. *Med Sci Sports Exerc* 1998;30:301-10.
30. Jette AM, Delitto A. Physical therapy treatment choices for musculoskeletal impairments. *Phys Ther* 1997;77:145-54.
31. Abenhaim L, Rossignol M, Valat J, et al. The role of activity in the therapeutic management of back pain: report of the International Paris Task Force on Back Pain. *Spine* 2000;25(suppl):1-35.
32. van Tulder MW, Malmivaara A, Esmail R, et al. Exercise therapy for low back pain: a systematic review within the framework of the Cochrane Collaboration Back Review Group. *Spine* 2000;25:2784-96.
33. Hides JA, Jull GA, Richardson CA. Long term effects of specific stabilizing exercises for first-episode low back pain. *Spine* 2001;26:E243-8.
34. 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;25:106-14.
35. Carey TS, Garrett JM, Jackman A, et al. The outcomes and costs of care for acute low back pain among patients seen by primary care practitioners, chiropractors, and orthopedic surgeons: the North Carolina Back Pain Project. *N Engl J Med* 1995;333:913-7.
36. Birkmeyer NJ, Weinstein JN. Medical versus surgical treatment for low back pain: evidence and clinical practice. *Eff Clin Pract* 1999;2:218-27.
37. Schechtman JM, Schroth WS, Verme D, et al. Randomized controlled trial of education and feedback for implementation of guidelines for acute low back pain. *J Gen Intern Med* 2003;18:773-80.
38. Dey P, Simpson CW, Collins SI, et al. Implementation of RCGP guidelines for acute low back pain: a cluster randomised controlled trial. *Br J Gen Pract* 2004;54:33-7.
39. Freeborn DK, Shye D, Mullooly JP, et al. Primary care physicians' use of lumbar spine imaging tests: effects of guidelines and practice pattern feedback. *J Gen Intern Med* 1997;12:619-25.
40. Barnett AJ, Underwood MR, Vickers MR. Effect of UK national guidelines on services to treat patients with acute low back pain: follow up questionnaire survey. *BMJ* 1999;318:919-20.
41. Schers H, Wensing M, Huijsmans Z, et al. Implementation barriers for general practice guidelines on low back pain: a qualitative study. *Spine* 2001;26:E348-53.
42. Borkan JM, Cherkin DC. An agenda for primary care research on low back pain. *Spine* 1996;21:2880-4.
43. Borkan JM, Koes B, Reis S, et al. A report from the second international forum for primary care research on low back pain: reexamining priorities. *Spine* 1998;23:1992-6.
44. Delitto A, Shulman AD, Rose SJ. On developing expert-based decision-support systems in physical therapy: the NIOSH Low Back Atlas. *Phys Ther* 1989;69:554-8.
45. Fritz JM, Whitman JM, Flynn TW, et al. Factors related to the inability of individuals with low back pain to improve with a spinal manipulation. *Phys Ther* 2004;84:173-90.
46. 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.
47. Hicks GE, Fritz JM, Delitto A, et al. 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.
48. Werneke M, Hart DL. Categorizing patients with occupational low back pain by use of the Quebec Task Force classification system versus pain pattern classification procedures: discriminant and predictive validity. *Phys Ther* 2004;84:243-54.
49. Werneke M, Hart DL. Centralization phenomenon as a prognostic factor for chronic low back pain and disability. *Spine* 2001;26:758-65.
50. Long A, Donelson R. Does it matter which exercise? A randomized trial of exercise for low back pain. *Spine* 2004;29:2593-602.
51. Hazard RG, Haugh LD, Green PA, et al. Chronic low back pain: the relationship between patient satisfaction and pain, impairment, and disability outcomes. *Spine* 1994;19:881-7.
52. Beurskens AJ, de Vet HC, Koke AJ, et al. Measuring the functional status of patients with low back pain: assessment of the quality of four disease-specific questionnaires. *Spine* 1995;20:1017-28.
53. Hiebert R, Skovron ML, Nordin M, et al. Work restrictions and outcome of nonspecific low back pain. *Spine* 2003;28:722-8.
54. Severijns R, Vlaeyen JW, van den Hout MA, et al. Pain catastrophizing predicts pain intensity, disability, and psychological distress independent of the level of physical impairment. *Clin J Pain* 2001;17:165-72.
55. Hagen KB, Hilde G, Jamtvedt G, et al. The Cochrane review of advice to stay active as a single treatment for low back pain and sciatica. *Spine* 2002;27:1736-41.
56. Von Korff M, Saunders K. The course of back pain in primary care. *Spine* 1996;21:2833-9.
57. Croft P, MacFarlane GJ, Papageorgiou AC, et al. Outcome of low back pain in general practice: a prospective study. *BMJ* 1998;316:1356-9.