Subgrouping Patients With Low Back Pain: Evolution of a Classification Approach to Physical Therapy

Low back pain (LBP) is the reason for seeking care in nearly 50% of all patients presenting to outpatient physical therapy clinics. As indicated by the Guide to Physical Therapist Practice, physical therapists employ a wide range of interventions in the management of patients with LBP, including manual physical therapy (ie, spinal manipulation), therapeutic exercise, traction, modalities, and functional training. Although a variety of interventions are accepted as standard of care for patients with LBP, high-quality evidence from randomized clinical trials has failed to offer conclusive support for most interventions. Over 1000 randomized clinical trials investigating the effectiveness of conservative and surgical interventions for the management of LBP have been reported in the literature. Despite this plenitude of research, the evidence remains contradictory and inconclusive for many interventions. One explanation offered for the lack of evidence for many common interventions relates to study designs with broad inclusion criteria, resulting in heterogeneous samples. Research on interventions for LBP has traditionally not incorporated a reality recognized by clinicians: that it is not reasonable to expect everyone with nonspecific LBP to benefit from any single treatment approach. It has been advocated that researchers can improve the power of their research by using methods to match subgroups of patients to interventions from which they are likely to benefit.

The lack of conclusive research evidence has provided clinicians with little information to guide decision making in the selection of interventions for individual patients, resulting in suboptimal outcomes and wide variations in practice patterns. Clinicians agree that LBP is a heterogeneous condition, but there is disagreement as to the most appropriate methods for classifying these patients to improve clinical outcomes. Traditionally the medical model has attempted to classify individuals based on a pathoanatomical source of symptoms; however, identifying relevant pathology in patients with LBP has proved elusive and is identified in less than 10% of cases. Therefore, attempting to identify a pathoanatomical source will infrequently be useful for guiding decision making, particularly for physical therapists. The Guide to Physical Therapist Practice recognizes that a primary goal of the diagnostic process is to classify patients based on clusters of signs and symptoms.
not presumed pathoanatomical causes. The Guide and other advocates of evidence-based practice also promote that effective sub grouping methods should ultimately direct decision making towards the most effective management strategies.\textsuperscript{106,107}

A good deal of work, both theoretical and experimental, has been performed by Physical Therapists describing subgroups of patients based on clusters of signs and symptoms and proposing a particular intervention strategy as most effective.\textsuperscript{13,91,105,115} One approach to classifying patients with LBP based on signs and symptoms was described by Delitto and colleagues in 1995.\textsuperscript{35} Research has supported that decision making based on this classification structure results in better outcomes for physical therapy than decision making based on alternative procedures.\textsuperscript{18,45} No system of patient management should be considered static, and it is necessary to incorporate new evidence into existing systems. Since the original proposal by Delitto et al,\textsuperscript{35} several studies have been conducted that expound on the specific signs and symptoms used to identify patient subgroups and the specific interventions and protocols that may be most beneficial for patients in a particular subgroup. The purpose of this clinical commentary is to describe the evidence published in recent years that may impact the classification system originally proposed by Delitto et al,\textsuperscript{35} and to discuss its implications for the physical therapy management of patients with LBP.

### TABLE 1

<table>
<thead>
<tr>
<th>Classification</th>
<th>Original Classification Criteria</th>
<th>Updated Classification Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manipulation</td>
<td>Asymmetrical lateral flexion ROM (ie, capsular pattern of motion restriction)</td>
<td>No symptoms distal to the knee</td>
</tr>
<tr>
<td></td>
<td>Unilateral LBP without symptoms into the lower extremities</td>
<td>Recent onset of symptoms (&lt;16 d)</td>
</tr>
<tr>
<td></td>
<td>Asymmetrical bony landmarks of the pelvis</td>
<td>Low FABQW score (&lt;19)</td>
</tr>
<tr>
<td></td>
<td>Positive sacroiliac dysfunction tests (ie, supine long sit test, prone knee bend test, standing flexion test)</td>
<td>Hypomobility of the lumbar spine</td>
</tr>
<tr>
<td>Stabilization</td>
<td>Frequent recurrent episodes of LBP with minimal perturbation</td>
<td>Younger age (&lt;40 y)</td>
</tr>
<tr>
<td></td>
<td>Hypermobility of the lumbar spine</td>
<td>Greater general flexibility (postpartum, average SLR ROM &gt;91°)</td>
</tr>
<tr>
<td></td>
<td>Previous history of lateral-shift deformity with alternating sides</td>
<td>“Instability catch” or aberrant movements during lumbar flexion/extension ROM</td>
</tr>
<tr>
<td></td>
<td>Frequent prior use of manipulation with dramatic but short-term results</td>
<td>Positive findings for the prone instability test</td>
</tr>
<tr>
<td></td>
<td>Trauma, pregnancy, or use of oral contraceptives</td>
<td>For patients who are postpartum:</td>
</tr>
<tr>
<td></td>
<td>Relief with immobilization (eg, bracing)</td>
<td>· Positive posterior pelvic pain provocation (P4), and ASLR and modified Trendelenburg tests</td>
</tr>
<tr>
<td>Specific exercise</td>
<td>Symptoms centralize with lumbar extension</td>
<td>Symptoms distal to the buttock</td>
</tr>
<tr>
<td>Extension</td>
<td>Symptoms peripheralize with lumbar flexion</td>
<td>Symptoms centralize with lumbar extension</td>
</tr>
<tr>
<td>Flexion</td>
<td>Symptoms centralize with lumbar flexion</td>
<td>Symptoms peripheralize with lumbar flexion</td>
</tr>
<tr>
<td></td>
<td>Symptoms peripheralize with lumbar extension</td>
<td>Directional preference for extension</td>
</tr>
<tr>
<td></td>
<td>Diagnosis of lumbar spinal stenosis</td>
<td>Older age (&gt;50 y)</td>
</tr>
<tr>
<td>Lateral shift</td>
<td>Visible frontal plane deviation of the shoulders relative to the pelvis</td>
<td>Directional preference for flexion</td>
</tr>
<tr>
<td></td>
<td>Asymmetrical side-bending active ROM</td>
<td>Imaging evidence of lumbar spinal stenosis</td>
</tr>
<tr>
<td></td>
<td>Painful and restricted extension active ROM</td>
<td>Visible frontal plane deviation of the shoulders relative to the pelvis</td>
</tr>
<tr>
<td>Traction</td>
<td>Signs and symptoms of nerve root compression</td>
<td>Signs and symptoms of nerve root compression</td>
</tr>
<tr>
<td></td>
<td>No movements centralize symptoms</td>
<td>No movements centralize symptoms</td>
</tr>
</tbody>
</table>

*Abbreviations: ASLR, active straight-leg raise; FABQW, Fear-Avoidance Beliefs Questionnaire-Work Subscale; LBP, low back pain; ROM, range of motion; SLR, straight-leg raise.*
OVERVIEW OF THE CLASSIFICATION SYSTEM

The classification system described by Delitto and colleagues was intended for patients with acute, or an acute exacerbation of, LBP causing substantial pain and limitations in daily activities. After screening patients for any medical red flags, the system proposed using the information gathered from the history and physical examination to place a patient into 1 of 4 basic classification categories: manipulation, specific exercise (flexion, extension, and lateral-shift patterns), stabilization, and traction. The signs and symptoms originally proposed as the criteria for placing a patient into one of these categories are listed in TABLE 1, and the intervention procedures originally proposed for each category are listed in TABLE 2. The system was based on clinical experience and the evidence available at the time. In the sections below we will review recent evidence that should inform the classification criteria and intervention procedures used for each category.

MANIPULATION CLASSIFICATION

Many randomized clinical trials have found spinal manipulation to be more effective than placebo or other interventions for patients with LBP. Conversely, other studies have shown that manipulation is not more effective than other treatments. The incongruous results of previous trials have led some to suggest that manipulation may be effective, but only for a subgroup of patients with LBP. Further consideration of recent evidence for examination and intervention procedures may help to clarify procedures to identify and manage patients in a manipulation subgroup.

Examination Considerations

Traditionally, classifying a patient as needing manipulation has relied heavily on mobility assessments and special tests based in biomechanical theories, and the examination procedures related to these theories were originally advocated as important classification criteria (TABLE 1). Many of these diagnostic tests have been found to have poor reliability and questionable validity and therefore no longer appear to be the preferred method for identifying patients needing manipulation. Recent research has focused on identifying baseline examination factors that are associated with benefiting from manipulation interventions without assumptions based on theory or tradition. Studies examining predictors of response to chiropractic treatment using manipulation have reported that patients with shorter duration of symptoms and the absence of leg pain are most likely to benefit.

We have pursued the development of a multivariate clinical prediction rule (CPR) to accurately identify patients who fit a manipulation classification. A CPR is a tool designed to assist the classification process and improve decision making by using evidence to determine which patients are likely to benefit from a specific treatment strategy. The goal of the CPR for the manipulation classification is to identify patients with LBP who are likely to respond to manipulation with rapid and sustained improvement. Flynn et al developed a CPR for the manipulation classification by examining predictors of improvement defined as a 50% or greater reduction in self-reported disability occurring over 2 treatment sessions in 71 patients with nonradicular LBP. The CPR included 5 factors: current symp-
tum duration of less than 16 days, a score on the work subscale of the Fear-Avoidance Beliefs Questionnaire (FABQ)\(^{195}\) of less than 19, hypomobility of the lumbar spine as assessed with posterior-to-anterior pressure, internal rotation of at least 1 hip greater than 35°, and symptoms not extending distal to the knee. When 4 of these 5 factors were present, patients were highly likely to improve (positive likelihood ratio [LR], 24), while the presence of 2 or fewer factors was almost always associated with a failure to improve (negative LR, 0.09). To put these results in perspective, if it is assumed that about 50% of all patients with nonradicular LBP would improve with manipulation, the likelihood of improvement would increase to 97% when at least 4 factors were present and decrease to 9% when 2 or fewer factors were present.

A follow-up study\(^{25}\) was carried out to examine the validity of the CPR by randomly assigning 131 patients to receive a standardized exercise program with or without manipulation and by examining the results in subgroups of patients based on their status on the manipulation classification CPR. The results demonstrated that patients who were positive on the CPR (ie, 4 or more factors) and received manipulation experienced greater improvement in pain and disability in short-term (at 1 and 4 weeks) and long-term (6 months) follow-ups than patients who were negative on the CPR (ie, fewer than 4 factors) and received manipulation (FIGURE 1). Patients who were positive on the CPR and received manipulation also experienced greater short- and long-term improvements in pain and disability than patients who were positive on the CPR but received the exercise intervention. These results indicate that the subgroup of patients identified by the CPR is uniquely responsive to a manipulation intervention.

The criteria for identifying patients in the manipulation classification have evolved from factors based largely on biomechanical theory to factors identified through prospective analysis with comparisons to clinical outcomes. Studies in this area appear to consistently support 2 factors (short duration of symptoms and no leg pain) as important criteria for the manipulation classification\(^{4,44,100}\) and the presence of at least 4 of the 5 CPR factors increases accuracy of predicting success even further. The value of a classification approach is not only the ability to identify the patients likely to benefit from a particular intervention, but also the ability to identify patients who need a different approach. Patients with 2 or fewer CPR factors appear very unlikely to improve with manipulation and likely need an alternative intervention. It is also important to note that patients over the age of 60 or with signs of nerve root compression were excluded from consideration in the studies developing this CPR, as were patients with diagnoses of spondylolisthesis, osteoporosis, or any concerns of bony abnormality or weakness. Manipulation is generally considered to be contraindicated in these subgroups,\(^{42,102}\) although some believe that manipulation may be appropriate for at least some patients with signs of nerve root compression.\(^{19,108}\)

**Management Considerations**

Biomechanical theories traditionally used to identify patients for the manipulation classification have also supported the need for precise techniques to address specific dysfunctions.\(^{72,104}\) The importance of the choice of a specific manipulation technique has recently been challenged as traditional theories underlying manipulation are questioned.\(^{26,29}\) Although evidence is sparse, a few studies have found greater benefit from thrust manipulation techniques versus nonthrust mobilization for the lumbosacral region.\(^{19,92}\) Although manipulation is generally recommended as superior to mobilization procedures,\(^{29}\) there is presently no evidence for the superiority of one manipulation technique over another.\(^{29}\) It is possible that the choice of a specific manipulation technique may not be as important as previously thought.\(^{29}\)

Originally the manipulation classification proposed by Delitto and colleagues\(^{35}\) incorporated traditional biomechanical approaches to technique selection, distinguishing techniques directed towards the sacroiliac or lumbar region (**TABLE 2**). Recent evidence, however, suggests that the effects of manipulation may not be as specific as once believed. For example, Beffa et al\(^{9}\) examined the relationship between manipulation targeted to specific spinal levels and the spinal levels actually producing a cavitation during the technique. The authors found no correlation between the spinal levels producing cavitation sounds and the levels targeted by the technique. Haas et al\(^{20}\) examined short-term outcomes of patients with neck pain randomized to receive manipulation targeted to spinal segments thought to have increased stiffness based on clinical examination or targeted to randomly selected segments, and found no differences in patient-reported pain or stiffness. Kent et al\(^{25}\) systematically reviewed the evidence on the effect of the discretion given to clinicians to choose techniques for a particular patient on outcomes in randomized trials examining manual therapy and found that although the evidence was limited, there was no suggestion that allowing clinicians to select techniques for patients improved outcomes compared with studies using predefined manipulation protocols.\(^{75}\)
cumulating evidence suggests that the most important factor to achieve optimal outcomes with manipulation may be the accurate identification of patients who are likely to respond rather than the selection of specific techniques.

### STABILIZATION CLASSIFICATION

The concept of a subgroup of patients with LBP related to spinal instability has been described for decades, but was initially discussed as a mechanical condition of excessive movement between adjacent vertebrae that required immobilization or surgical stabilization.23,96,109 The original classification system proposed in 199523 reflected this perspective, labeling this subgroup “immobilization” and recommending examination criteria and interventions designed to manage patients with excessive segmental movement (TABLES 1 and 2). Recent research has provided a somewhat different perspective by emphasizing the importance of spinal muscles in maintaining and restoring spinal stability, shifting the focus of rehabilitation from immobilization to stabilization.23,32,69,70 In the last few years, this research has greatly increased the popularity of exercise interventions designed to enhance the stabilizing capacity of spinal muscles.109 There have been several randomized trials published to investigate the effectiveness of lumbar stabilization exercises for patients with LBP that have reported inconsistent results.23,56,67,85,99,109 As previously suggested, these conflicting results may suggest that stabilization exercises are effective for some, but not all, patients with LBP. Further evaluation of recent evidence on the examination and intervention procedures related to the subgroup of patients most likely to benefit from stabilization exercise may improve identification and management of these patients.

**Examination Considerations**

Delitto et al35 originally described the classification criteria for a stabilization subgroup that focused on identifying patients presumed to have excessive segmental movements of the spine (TABLE 1), such as recurrent LBP episodes, frequent manipulation or self-manipulation with short-term relief, trauma, pregnancy, oral contraceptive use, and positive response to immobilization of the spine. Recent surveys of physical therapists suggest that this perspective on identifying patients for stabilization interventions remains prevalent.33,75

Most research conducted to identify stabilization classification criteria has examined the usefulness of clinical examination findings for identifying radiographic evidence of excessive motion between vertebrae.1,40,51 However, the validity of this approach has been questioned based on studies showing wide interindividual and intrindividual variations in spinal motion characteristics in asymptomatic subjects, making it difficult to establish thresholds identifying a spine as unstable.14,63

Using the amount of segmental motion as the standard against which examination variables are judged also fails to account for the important role of the spinal muscles,46 and it is inconsistent with the goal of a classification approach. Classification seeks to identify patients likely to respond to a specific treatment approach, not those with a particular imaging finding.

We have sought to identify examination criteria for the stabilization classification by developing a CPR for this subgroup. Hicks et al65 provided 8 weeks of stabilization training targeting the multifidus/erector spinae, transversus abdominus, and oblique abdominal

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>Special Tests Suggested to Be Important Examination Criteria for Identifying Patients in the Stabilization Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examination</td>
<td>Description</td>
</tr>
<tr>
<td>Prone instability test&lt;sup&gt;46&lt;/sup&gt;</td>
<td>The patient lies prone with the body on an examining table and legs over the edge with feet resting on the floor. While the patient rests in this position, the therapist applies posterior-to-anterior pressure to the lumbar spine. Any provocation of pain is reported. Then the patient lifts the legs off the floor 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</td>
</tr>
<tr>
<td>Posterior pelvic pain provocation (P4) test&lt;sup&gt;97&lt;/sup&gt;</td>
<td>The patient is supine. The therapist passively flexes the patient’s hip to 90° and applies a posteriorly directed force through the longitudinal axis of the femur. The test is positive if the patient reports a deep pain in the gluteal area during the test</td>
</tr>
<tr>
<td>Active straight-leg raise test&lt;sup&gt;87&lt;/sup&gt;</td>
<td>The patient is supine with straight legs and feet 20 cm apart. The patient is instructed to lift the legs one after the other approximately 20 cm above the table without bending the knee. The patient is asked to score the difficulty of the task on a 6-point scale (0, no difficulty at all; 1, minimally difficult; 2, somewhat difficult; 3, fairly difficult; 4, very difficult; 5, unable to do). Any score greater than 0 is a positive test</td>
</tr>
<tr>
<td>Provocation of the long dorsal sacroiliac ligament&lt;sup&gt;118&lt;/sup&gt;</td>
<td>The patient is supine. The therapist palpates the long dorsal sacroiliac ligament bilaterally. A positive test occurs if at least 1 side is painful, and the pain persists at least 5 seconds after the removal of the therapist’s hand</td>
</tr>
<tr>
<td>Provocation of the pubic symphysis with palpation&lt;sup&gt;4&lt;/sup&gt;</td>
<td>With the patient in supine the entire front side of the pubic symphysis is palpated gently. If the palpation causes pain that persists more than 5 seconds after the removal of the therapist’s hand, it is recorded as positive</td>
</tr>
<tr>
<td>Modified Trendelenburg test&lt;sup&gt;4&lt;/sup&gt;</td>
<td>The therapist is behind the standing patient. The patient is asked to stand on one foot while flexing the opposite knee and hip to 90°. The test is positive if the hip descends on the flexed side</td>
</tr>
</tbody>
</table>
muscules to 54 patients with nonradicular LBP. Using a definition of improvement (50% reduction in self-reported disability), the authors identified 4 factors that were predictive of improvement: age less than 40 years, average straight-leg raise (SLR) range of motion (ROM) greater than 91°, aberrant movements during sagittal plane lumbar ROM, and a positive prone instability test (TABLE 3). A preliminary CPR was defined as positive when 3 or more of these factors were present; however, the predictive accuracy of the stabilization CPR (positive LR, 4.0) was not as strong as the manipulation CPR. Assuming that a patient has a 50% chance of improving with a stabilization intervention, a positive CPR increases the probability to 80%. Greater accuracy was found for identifying patients who were not likely to receive even minimal benefit (5 or fewer points of improvement on the Oswestry) from a stabilization intervention. Four factors predictive of failure included a negative prone instability test, absence of aberrant movements during sagittal plane lumbar ROM, absence of lumbar hypermobility (assessed with posterior-to-anterior pressure), and a score of less than 9 on the FABQ physical activity subscale. The presence of at least 3 of these findings was highly predictive of failure (positive LR, 18.8), indicating that if a patient was presumed to have a 25% probability of failing, the presence of at least 3 of these factors would increase the probability of failure to 86%.

Stuge and colleagues have proposed additional factors to identify some women with posterior pelvic girdle pain who are postpartum as likely to benefit from stabilization treatment. The criteria used to define this subgroup are women who are postpartum with buttock pain and a composite of positive tests: posterior pelvic pain provocation (P4) test, active straight-leg raise (ASLR) test, provocative of the long dorsal sacroiliac ligament, provocation of the pubic symphysis with palpation, and the modified Trendelenburg test4 (TABLE 3).

The variables identified in these studies are generally consistent with current theories emphasizing the importance of spinal muscles as a component of stabilization. Patients in the stabilization classification appear to be those who are generally flexible (ie, younger, excessive SLR ROM) or with increased flexibility (ie, postpartum), possibly with increased segmental spinal movement (ie, hypermobility), whose spinal muscles do not provide adequate stabilization (ie, aberrant movements, and positive prone instability, ASLR, and modified Trendelenburg tests). Further research is necessary to refine and validate the criteria defining the stabilization classification.

Management Considerations

The original classification system proposed interventions focused on restricting movement that was presumed to be excessive for patients in a stabilization classification. Recommendations included avoiding end-range positions of the spine and bracing for more severe cases, along with spinal muscle strengthening exercises. Research on the stabilizing role of spinal muscles has shifted the focus of treatment for patients in the stabilization classification from avoiding to controlling movement. In particular, recent research has stressed the importance of the deep muscles of the spine for stabilization (ie, transversus abdominus, multifidus). This research has increased attention on stabilization exercise programs that emphasize specific retraining of these muscles. Others have focused stabilization exercise regimens on improving the strength and endurance of larger spinal muscles (ie, erector spinae, oblique abdominals, quadratus lumborum), creating some disagreement concerning optimal intervention strategies for patients in the stabilization classification.

Support for the specific-muscle approach to stabilization comes from randomized trials that have found better outcomes resulting from stabilization exercise programs centered on retraining appropriate activation of the transversus abdominus and/or multifidus muscles when compared to no treatment or multimodal treatment programs not explicitly focused on strengthening exercises.

Two recent studies have questioned if specific muscle retraining is the most effective approach to stabilization. Cairns et al randomized 97 patients with a prior history of LBP to specific muscle retraining or conventional physical therapy. Both groups received individually tailored exercise and manual therapy interventions. The specific muscle retraining group received additional instruction in retraining the multifidus and transversus abdominus, supplemented with written instructions and real-time ultrasound biofeedback as needed. No differences in outcome were detected after 12 weeks of treatment or at 1-year follow-up. Koumantakis et al also examined patients with recurrent LBP, randomizing 67 subjects to a specific-retraining group that focused initially on retraining the multifidus and transversus abdominus or to a general-strengthening group that concentrated on strengthening the large muscle groups of the spine (erector spinae, oblique abdominals). The authors found somewhat superior outcomes for the general-strengthening group following the 8-week treatment program, with no differences at the 20-week follow-up.

Further research is needed to identify the most effective intervention strategies for patients in the stabilization classification. Although many experts advocate the necessity of specifically retraining the deep spinal muscles, the evidence does not clearly support this approach. It appears that specific muscle retraining protocols are superior to treatments that do not include a well-defined strengthening component, but the superiority of a specific approach to muscle retraining over an approach that stresses general strengthening of the larger spinal muscles has not been demonstrated.
The existence of subgroups of patients who preferentially respond to repeated end-range movements was popularized by McKenzie several decades ago. Consistent with principles proposed by McKenzie, Delitto and colleagues identified a classification of patients for whom repeated exercises in a specific direction (flexion, extension, or a lateral shift) were proposed to be the appropriate intervention. The presence of the centralization phenomenon was the primary examination criterion proposed for membership in a specific-exercise classification, and the movement producing centralization determined the specific direction of exercise required for the patient. The first generation of randomized trials examining specific-exercise interventions found no evidence of benefit in trials examining specific-exercise interventions. A recent study used centralization as an inclusion criterion and examined the effectiveness of an extension specific-exercise protocol compared to a stabilization approach. The results showed better outcomes in the group receiving the extension protocol in this sample of patients who demonstrated centralization with extension movements at baseline. This is the first study to provide some evidence of the usefulness of centralization as a classification criterion for specific-exercise classifications.

An examination finding related to centralization that has also been studied as a classification criterion for specific exercise is the finding of a directional preference. A directional preference is defined as a situation in which movement in one direction improves pain and limitation of ROM, and movement in the opposite direction causes signs and symptoms to worsen. A patient who exhibits centralization with a movement would be considered to have a directional preference for that movement; but centralization is not required, making directional preference a broader category of patients. Long et al studied patients with a directional preference, randomizing them to receive a specific-exercise intervention in the direction that matched their directional preference, a specific-exercise intervention in an unmatched direction, or a control group. The results indicated greater reductions in pain over a 2-week follow-up period when the specific-exercise regimen was matched to the patient’s directional preference as compared to the group receiving the unmatched-exercise direction. Additional research is needed to examine the usefulness of centralization and directional preference for identifying patients likely to respond to specific-exercise interventions. Future research may also identify additional examination criteria for specific-exercise classifications.

Management Considerations

The basic premise advocated for treating patients in a specific-exercise classification is to use repeated end-range movements in the direction that caused centralization. This approach was recommended in the original classification system, leading to 3 categories based on the centralizing movement (flexion, extension, or a lateral shift). Two recent systematic reviews have data from 6 randomized or quasi-experimental studies investigating the effects of treatment provided according to principles proposed by McKenzie, a large component of which is repeated end-range movement in the direction of centralization. These reviews found greater reductions in pain and disability for treatments based on McKenzie principles in the short term, but the differences were small in magnitude and no longer significant at long-term follow-up. Studies included in these reviews used broad inclusion criteria, which may explain the small treatment effects. The reviews also included only studies with treatments provided according to McKenzie principles. Examining a broader group of studies may provide additional insight into the management of specific-exercise classifications.

The most common direction used with patients in a specific-exercise classification is extension, and extension protocols have been studied the most. The study by Long et al included 230 patients with LBP and/or sciatica who had a directional preference, and randomly assigned them to receive exercises matching their preference, exercises opposite the identified preference, or a control group. For 83% of the patients extension was the direction of preference. The matched-direction treatment protocol in this study included 2 components: repeated end-range exercises (eg, prone press-ups) and patient education. Although patients with an extension preference were not considered separately, the predominance of an extension preference makes it likely that the matched direction treatment was more effective for the subgroup.
et al\textsuperscript{35} studied 260 patients with chronic LBP with or without sciatica, comparing an extension-oriented protocol with a general-strengthening program. In this study, the extension protocol included repeated end-range extension exercise along with mobilization performed by a physical therapist. Although the sample was heterogeneous, short-term results favored the extension protocol group, but the treatment effects were small.\textsuperscript{100} Browder et al.\textsuperscript{21} in a sample of patients who centralized with extension, also found better results for a group receiving mobilization (graded mobilization to promote extension) along with extension exercises and patient education. The optimal intervention strategy for patients in the extension specific-exercise classification may be a combination of exercise and mobilization to promote end-range extension.

Flexion specific-exercise classification appears to be less common\textsuperscript{83} and most likely occurs in patients who are older, often with a medical diagnosis of lumbar spinal stenosis.\textsuperscript{47} Interventions originally advocated for patients in the flexion specific-exercise classification were flexion-oriented exercises (eg, knee-to-chest, pelvic tilts, etc), and traction with the patient in a position of spinal flexion if there was a diagnosis of lumbar spinal stenosis.\textsuperscript{25} Little research has been performed examining the effectiveness of intervention strategies for these patients, and most research has focused on patients with stenosis instead of a more general flexion specific-exercise classification. Case studies of patients with stenosis have advocated intervention strategies, including mobilization or manipulation for the lumbar spine and/or hip, general lower extremity strengthening, neural mobilizations, and a walking program possibly facilitated with body weight-supported treadmill ambulation.\textsuperscript{49,55,123} A recent randomized trial\textsuperscript{122} examined patients over age 50 with a directional preference for flexion and imaging evidence of lumbar spinal stenosis. One group received manual therapy (mobilization or manipulation of the spine and/or lower extremity), exercise to address impairments of strength or flexibility, and a body weight-supported treadmill-walking program. The other group received flexion-oriented exercises, a treadmill-walking program (without body weight support), and subtherapeutic ultrasound. Better outcomes were reported by the group receiving manual therapy, exercise, and body weight-supported walking.\textsuperscript{122} The multimodal intervention protocol precludes conclusions on any individual procedure; however, the results suggest that interventions for patients in the flexion specific-exercise classification should include several components other than flexion-oriented exercise.

The third movement direction in the specific-exercise classification is a lateral shift, which is considerably less common than flexion or extension categories.\textsuperscript{50,83} For example, only 7% of the subjects with a directional preference studied by Long et al.\textsuperscript{83} had a preference for a lateral shift movement. In the original classification system, treatment for patients in the lateral shift specific-exercise classification included repeated end-range lateral-shifting exercise or traction (mechanical or autotraction).\textsuperscript{35,82} Harrison et al.\textsuperscript{46} reported the results of a nonrandomized comparison of patients with a visible lateral shift who received a program of repeated lateral-shift exercises and mechanical traction, and reported greater pain reductions and correction of the shift, compared to a group of patients receiving no treatment. Gillan et al.\textsuperscript{41} studied 40 patients with a visible lateral shift, randomizing patients to management with repeated end-range lateral-shift exercises or nonspecific advice and massage. The group receiving the lateral-shift exercises experienced more rapid resolution of the lateral shift, but no differences were found in disability outcomes after 3 months.\textsuperscript{41} Further research is required to clarify the most effective intervention strategies for patients in the lateral-shift specific-exercise classification.

### Traction Classification

Although there was no evidence to support the contention, Dellito and colleagues\textsuperscript{35} hypothesized that there is a subset of patients with LBP who would likely benefit from traction. The examination criteria defining this subgroup was proposed to be the presence of lower extremity symptoms and signs of nerve root compression and the absence of centralization with movement testing. There continues to be a lack of evidence supporting the use of traction for patients with LBP, and the intervention is generally not recommended by systematic reviews and practice guidelines.\textsuperscript{28,70,117}

Studies that have shown no benefit from using traction have not sought to identify the patients who are most likely to benefit from the intervention, but have instead used nonspecific inclusion criteria, essentially allowing all patients fitting a broad definition of acute or chronic LBP to enter.\textsuperscript{31,121} Recent systematic reviews on the effectiveness of traction as an intervention for patients with LBP,\textsuperscript{24,61} while acknowledging the lack of any evidence to support the use of traction, also note that this may be related to the fact that studies have included “mixed groups” of patients rather than homogenous samples presumed to be likely to benefit from the intervention.

Similar to the recommendations of Delitto et al.\textsuperscript{35} the most common examination criterion cited by clinicians as an indication for traction is the presence of signs of nerve root compression.\textsuperscript{61} Buerskens et al.\textsuperscript{10} compared the effects of mechanical traction (maximum force, 35%-50% of body weight) to sham traction (maximum force, 20% of body weight) for 12 sessions over 5 weeks in patients with nonspecific LBP of at least 6 weeks in duration. Following treatment, there was no difference between groups for perceived recovery. The authors performed a secondary analysis in an attempt to identify a subgroup of patients responding positively to traction and considered the following variables:
age, sex, duration of episode, radiation of symptoms below the knee, general health, severity of symptoms, maximum traction force used, and the physical therapist’s belief that traction would be beneficial for the patient. None of the aforementioned subgroups were found to have experienced a greater benefit with mechanical traction as compared to sham traction.10 The authors did not investigate all examination criteria proposed in the original classification system, and perhaps factors such as signs of nerve root compression and absence of centralization will prove to be important examination criteria for identifying a traction classification.

We believe the available research can be interpreted to indicate that the majority of patients with LBP are not appropriate for a traction intervention and, therefore, traction should not be widely used for patients with LBP. It does not appear that current clinical decision making used by physical therapists is adequate for identifying which patients with LBP may respond to a traction intervention.10 Future research is needed to determine if examination criteria exist that can identify a patient who is likely to respond to traction. Additional research is also necessary to define the parameters that may maximize any treatment effect (eg, traction force and duration, patient position, etc).

**FURTHER CONSIDERATIONS**

In 1998, Riddle105 provided a review and critique of classification systems for the management of patients with LBP, including the system proposed by Delitto and colleagues22 using defined LBP, including the system proposed by areas have been addressed through ongoing research.

One deficient area105 was the lack of specific, reliable criteria for inclusion into each classification. Further research identifying examination criteria for the manipulation, stabilization, and specific-exercise groups has been conducted with distinct criteria identified for each classification. Interrater reliability of the individual factors identified for the manipulation,91 stabilization,23,66 and specific-exercise68 subgroups has been published. The reliability of classification judgments made using the system was also an area of concern105 that has now been examined in several studies. Heiss et al64 studied the reliability of the classification system among 4 different raters who were inexperienced with using the system. Following a 1-day training session, the clinicians classified 45 consecutive patients with LBP, with each rater blind to the others’ decisions. Three out of 4 rater pairs achieved a kappa value of 0.45 (55% agreement). This kappa value was slightly lower than that reported by Fritz and George104 (65% agreement with

---

**TABLE 1.** Classification decision-making algorithm. Abbreviations: FABQ, Fear-Avoidance Beliefs Questionnaire; FABQPA, FABQ Physical Activity Subscale; FABQW, FABQ Work Subscale; LBP, low back pain; ROM, range of motion; SLR, straight-leg raise. Adapted with permission from Fritz et al.43

---

**TABLE 2.** Classification decision-making algorithm. Abbreviations: FABQ, Fear-Avoidance Beliefs Questionnaire; FABQPA, FABQ Physical Activity Subscale; FABQW, FABQ Work Subscale; LBP, low back pain; ROM, range of motion; SLR, straight-leg raise. Adapted with permission from Fritz et al.43
a kappa value of 0.56) in a study using more experienced examiners. The classification system has continued to evolve, and a recent study examined the reliability of a more explicit decision-making algorithm (FIGURE 2), with the traction classification removed and using therapists with varying levels of experience with the system.43 The overall agreement between therapists was 76%, with a kappa value of 0.60 (95% CI: 0.56, 0.64). No differences in agreement existed based on experience.43

Additional criteria for a classification system are that it should be simple, easy to understand, and indicate if special training is required.22,105 While the original algorithm for the classification system was quite complex, with multiple steps and considerations, modifications made based on emerging evidence has simplified the decision-making scheme (FIGURE 2), which appears to have improved the reliability of the system and should increase the ability to incorporate decision-making into clinical practice without specific training. It also appears that the intervention strategies proposed by the classification system can be applied effectively by physical therapists regardless of clinical experience. Whitman et al22 found no difference in outcomes associated with therapists' years of experience in a group of patients with LBP who received manipulation or stabilization exercise interventions.

While it is useful to have evidence for the validity of the specific interventions in each classification, perhaps the most important factor to consider is whether overall outcomes are improved when the system is used as compared to some alternative approach.22 Two randomized trials19,45 have compared use of this classification system to other decision-making approaches for the management of patients with LBP in physical therapy. Fritz et al45 randomly assigned 78 patients with acute, work-related LBP to treatment based on the classification system or a current clinical practice guideline.12 All patients attended a mean of 5 physical therapy sessions. At the 4-week follow-up, patients treated with the classification approach exhibited significantly greater improvement in disability and general health status, higher satisfaction, and increased likelihood of returning to work than patients treated based on the guidelines. More recently Brennan and colleagues46 randomly assigned 123 patients to receive treatment according to the stabilization, manipulation, or specific-exercise classification, then compared patients matched or unmatched to their treatment group. At the 4-week and 1-year follow-ups, patients receiving matched treatment exhibited significantly greater reductions in disability than those in the unmatched-treatment group.46 Both studies provide support for the classification system as a decision-making scheme to place patients with LBP into subgroups that indicate the interventions that are most likely to provide benefit.

CONCLUSION

In 1989, Rose106 hypothesized that a useful classification system for the management of patients with LBP should lead to the identification of specific subgroups from data collected during the initial history and physical examination, which in turn guided the selection of optimal intervention strategies. Evidence that has emerged since the proposal of this hypothesis has confirmed its prescience. It should be recognized that the process of developing a classification system is dynamic and it is likely that further modifications will inevitably be made. However, it does appear that the outcomes of physical therapy care can be improved when patients are classified and treated accordingly.

REFERENCES


Cleland JA, Fritz JM, Childs JD, Kulig K. Comparision of the effectiveness of three manual physical therapy techniques in a subgroup of patients with low back pain who satisfy a clinical prediction rule: study protocol of a randomized clinical trial [NCT00257998]. *BMC Musculoskelet Disord.* 2006;7:11.


CORRECTION

In 2005, Hicks et al.² published criteria related to the prediction of clinical success or failure for patients with low back pain, who were undergoing a program of trunk strengthening and stabilization exercises. One of the criteria related to the prediction of failure with this treatment approach was a low score (<9) on the Physical Activity Subscale of the Fear-Avoidance Beliefs Questionnaire (FABQPA).³ This criterion was erroneously published by the JOSPT as “a score of 9 or higher” in the clinical commentary from Fritz and colleagues.¹ The error occurs in the text on page 295 of the commentary. The criterion is correctly cited in FIGURE 2 on page 298.

We apologize for this error and have corrected the article by Fritz et al., which is available to members and subscribers for download on the JOSPT web site (www.jospt.org). *

REFERENCES

1. Fritz JM, Cleland JA, Childs JD. Subgrouping patients with low back pain:
