Efficacy of Flexion and Extension Treatments Incorporating Braces for Low-Back Pain Patients With Retrodisplacement, Spondylolisthesis, or Normal Sagittal Translation

Kevin F. Spratt, PhD,* James N. Weinstein, DO,* Thomas R. Lehmann, MD,t Joyce Woody,* and Hutha Sayre, RN*

Radiographic instability seemingly enjoys the status of a well-defined clinical syndrome. The concept is widely used, and specific treatments, usually spinal fusion, are routinely performed based on the diagnosis. The minimum standards necessary to establish radiographic instability as a legitimate clinical syndrome have not been established, however. The primary purpose of this study was to determine if treatment invoking bracing, exercise, and education controlling either flexion or extension postures, would result in a distinctive pattern of favorable or unfavorable results, depending on the type of radiographic instability (retrodisplacement or spondylolisthesis). Fifty-six patients meeting strict study inclusion and radiographic evaluation criteria were assigned to a bracing treatment (flexion, extension, placebo-control) according to a randomization scheme, designed to ensure equal representation of translation categories (retro, normal, spondy) across treatment groups, and assessed at admission and 1-month follow-up. The sample was relatively evenly divided between men (46%) and women (54%), and by age. Translation classification was related to both gender and age, with men more likely classified as retro and women more likely spondy, and patients in their 20's having lower incidence of spondy and higher incidence of normal translation. Translation classification was not related to selected indices of low-back pain history. Brace treatments were not shown to reduce patient range of motion or lessen trunk strength. A significant treatment by time interaction for the modified pain interference (VAS) scale indicated improvement for patients in extension compared with patients in flexion and control-placebo treatments. In conjunction with no significant three-way interaction between treatment, translation classification, and time, it was hypothesized that radiographic instability might more appropriately be considered a corroborative sign of advanced discogenic problems. Improvement in extension treatment, regardless of the type of radiographic abnormality, suggests that the treating clinician might consider extension treatment for chronic low-back pain patients.

The precise diagnosis is unknown in 80-90% of patients presenting with disabling low-back pain. In a population study, Valkenburg and Haanen screened more than 6,000 men and women and were able to document objective evidence of disc prolapse or lumbago in fewer than 10% of their sample. In a general practice study, Dillane et al found that 79% of men and 89% of women presenting with a first episode of low-back pain were classified as having low-back pain of unknown origin. Nachemson et al have estimated that only 15% of chronic low-back pain cases have some demonstrated patho-anatomic explanation.

From the *Spine Diagnostic and Treatment Center, Department of Orthopaedic Surgery, University of Iowa Hospitals and Clinics, Iowa City, Iowa and the +Louisville Orthopedic Clinic, Louisville, Kentucky. Supported in part by NIH grant #AR 34344-03 and Camp International, Inc (J. Weinstein, principle investigator). Accepted for publication May 11, 1992.
many epidemiology studies. Lumbar fusions have been and continue to be performed because of this radiographic abnormality when accompanied by chronic pain, despite the fact that there is no universally accepted definition of instability. Even when surgery is not indicated, the diagnosis of instability may effect disability awards. Therefore, establishing an objective definition of instability may be essential to conducting meaningful epidemiologic studies and clinical trials as well as to the monitoring of clinical practice related to low-back pain.

Biomechanically, the definition of instability depends on the determination of abnormal motion to applied loads.30 Because there are three axes of rotation and three planes of translation for movement to occur, there is potential for several types of instability or observable excessive motions. Historically, interest and clinical treatment has focused on excessive translation in the sagittal plane.

There are no conclusive clinical data to support a cause-and-effect relationship between a radiographic finding of excessive translation in the sagittal plane (instability) and a low-back pain syndrome. Other than radiographic evidence, no set of objective clinical criteria are available to corroborate a diagnosis of radiographic instability. Given the recent evidence suggesting the potential for misclassifying radiographic translations in the sagittal plane,34 it seems especially important to develop corroborative clinical evidence associated with radiographic instability.

The necessary conditions for establishing a low-back pain syndrome have been well documented10,44:

1. The incidence of the abnormality should be higher in a patient population than in a normal population.
2. A reproducible pattern of historical or physical examination findings should characterize the low-back pain syndrome.
3. Stimuli specifically designed to exacerbate the condition should result in aggravation of symptoms or treatment specifically designed to aid the condition should result in alleviation of or palliation of symptoms.

Regarding the incidence of the radiographic abnormality, Knutsson17 described excessive displacement in 34% (48/140) of his clinical sample, and an additional 8% (11/140) were classified as spondylolytic spondylolisthesis. Morgan and King25 classified 29% (143/500) of patients with chronic low-back pain as having instability, with 10% (50/500) having spondylolisthesis. Schmorl and Junghanns,32 in cadaver studies, determined the incidence of retrolisthesis as 1.5% (9/600) and degenerative spondylolisthesis as 2.3% (14/600).32 Roentgenograms of these specimens were not assessed in flexion and extension, however. Friberg and Hirsch placed cadaver spines in flexion and extension and determined that 15% (15/100) of the specimens demonstrated roentgenographic instability. In their clinical cases, however, the incidence of instability was 27% (578/2104). Thus, this body of evidence suggests that the incidence of radiographic abnormality is greater in a low-back pain population relative to asymptomatic or general populations.

There are reports in the literature suggesting there are clinical signs corroborating the radiographic finding of excessive translations.2,20,25 A reproducible pattern of pain or symptom aggravation or relief in response to mechanical maneuvers or general activities of daily living has not been established, however. Work attempting to establish such links is underway but is not the focus of this report.

There are several reports of patients who have undergone lumbar fusion in response to radiographic findings of instability with favorable outcome.4,9,14,22,25 These results were from uncontrolled clinical trials, however, and thus provide limited validity for the hypothesis that instability represents a clinical syndrome. Because fusion is seen by many as the obvious and perhaps best treatment for instability because it surgically addresses the inherent abnormality, a prospective controlled clinical trial might be considered an excellent approach for establishing instability as a clinical syndrome. Given the nature of open surgery, however, application of fusion to necessary control groups, such as low-back pain patients without radiographic signs, is difficult.

The primary purpose of this study was to determine if a nonoperative treatment, involving bracing, exercise, and education controlling either flexion or extension postures, would result in a distinctive pattern of favorable or unfavorable results, depending on the type of sagittal translation abnormality. Specifically, it was hypothesized that patients with retrolisthesis, spondylolisthesis, and “normal” translation would respond differentially to treatments associated with flexion and extension posture. A secondary goal was to establish more objective criteria for prescribing braces for low-back pain patients.

Materials and Methods

Materials. A series of forms were developed to evaluate the patient at admission and at 1-month follow-up. Specifically, forms were developed that detailed the patient’s general demographics and medical history; low-back pain history; radiographic evaluation summary; a physical examination,35 including nonorganic signs42; and measures of trunk strength and range of motion.

The general demographics, medical history, and physical examination at initial presentation were used to establish general admission criteria to the study. The low-back pain history and evaluation of patient trunk strength and range of motion provided the outcome criteria for the study. Trunk strength and spine range of motion were evaluated using standard protocols.19

The primary outcome measure was a self-report version of
the pain interference scale developed by Million et al.24 and often referred to in the literature as the “Visual Analog Scale” (VAS). This scale was developed as an outcome measure for use in clinical trial studies and, in fact, was used initially to evaluate the efficacy of a corset treatment. The content of the 15 items was retained with slight modification and the instructions were modified to allow the scale to be completed by the patient. Each of the scales were made into lo-point visual analog scale. After completion, some items were reverse scaled so that high scores reflected greater wellness. The composite or total score was computed as the average of the 15 visual analog scale items. Thus, a patient’s score could range from 1 to 10, with larger scores indicating less pain interference, or greater wellness.

The original pain interference VAS was computed by “standardizing” each item as a z-score (mean, 0, and standard deviation, 1) and then summing these 15 z-scores to obtain a composite. Although the rationale for this procedure was presented in the original report, the benefits associated with allowing each item in the scale to contribute equally to the total was not seen as sufficient to outweigh the costs. Specifically, this approach forces the given study group to be its own norm group because a patient’s score simply reflects individual functioning relative to the specific group measured. In other words, a patient’s VAS score reflected a relative position in the group without consideration of the overall impairment status of that group. This approach makes it very difficult to use the VAS to compare results across studies. In addition, longitudinal assessments using this measure provide evidence about a patient’s relative improvement or worsening compared with the group, with no notion of a patient’s absolute improvement or worsening. As an example, consider a group of 10 patients assessed with the VAS at admission and at 1-month follow-up. Suppose that nine of the patient showed marked improvement and one patient stayed the same. The traditional VAS score for the patient who did not improve would be lower at follow-up, which might be interpreted as a sign of worsening. In other circumstances, a lower VAS score at follow-up could occur even if the patient showed improvement, but of a lesser magnitude than the cohort or a higher VAS score at follow-up could occur if a patient worsened, but less so than others in the cohort.

**Radiographic Criteria.** A frame designed to maximize flexion and extension postures while simultaneously avoiding obliquity from a true lateral projection was used to standardize the radiographic procedures used to obtain flexion and extension films for each patient. Two raters measured each flexion and extension film for relative translation in the sagittal plane at the L3-L4, L4-L5, and L5-S1 levels using the Lehmann method.18 described in detail by Shaffer et al.34

There are two commonly used methods for classifying radiographic instability, the maximum translation and the excursion approaches.1,13,38 Stokes and Frymoyer37 were unable to detect large amounts of anteroposterior excursion with biplanar radiographic techniques, nor able to correlate excursion with clinical symptoms and signs. Peacey and Shepherd29 also was unable to detect large amounts of excursion in patients with spondylolisthesis. In pilot testing leading to this study, relative angular motion at each motion segment (Hanley method)12 and excessive excursion (difference between anteroposterior translation measurements) did not correlate with clinical symptoms. When instability was defined as maximum displacement away from absolute vertical alignment on the flexion and extension films (i.e., either posterior [retrodisplacement] or anterior [spondylolisthesis]), interesting clinical features were observed. Specifically, retrodisplacement occurred more often in young men involved with lifting compared with spondylolisthesis, which occurred more often in older women, less involved with heavy lifting. Additionally, certain items, such as “leg pain with twisting”, were associated with spondylolisthesis. As a result of these findings, the maximum translation approach was adopted to classify patient instability. This definition is consistent with surgical practices in which patients generally receive lumbar fusion for the presence of spondylolisthesis rather than excessive mobility of a motion segment. In addition, the original works describing “instability” by Knutsson, l7 Morgan and King,25 and Friberg and Hirsch’ used a qualitative definition considering the relative alignment of the adjacent vertebral bodies (“retroposition”, “pseudospondylolisthesis”) rather than a quantitative definition related to excursion or hypermobility.

The rules for classifying translations as retrodisplacement (retrolisthesis), normal, or spondylolisthesis at each level are summarized below:

Classification Rules for Maximum Sagittal Translation (Not Excursion)

<table>
<thead>
<tr>
<th>Level</th>
<th>Retro Normal</th>
<th>Spondy</th>
</tr>
</thead>
<tbody>
<tr>
<td>L3-L4</td>
<td>≤3.5 mm</td>
<td>≥3.5 mm</td>
</tr>
<tr>
<td>L4-L5</td>
<td>≤3.5 mm</td>
<td>≥3.5 mm</td>
</tr>
<tr>
<td>L5-S1</td>
<td>≤3.5 mm</td>
<td>≥3.5 mm</td>
</tr>
</tbody>
</table>

Valid translation categorization required both raters to agree in their classification. Thus, if one rater classified a patient at normal and the other rater classified the patient as outside the normal range, then classification at that level for that view was coded as ambiguous, and patient inclusion and overall classification became dependent on classifications at other levels or from the other view of that level. In addition, if a translation was outside the normal range, but not in the retro or spondy range (e.g., 2.5 mm of translation), that classification was coded as ambiguous. Although these rules were considered quite strict, it was decided that such strictness was necessary to minimize the chances of erroneous patient classification. More recent work in the laboratory has demonstrated that a more conservative cut point may be necessary (i.e., displacement greater than 6 mm), especially when films of inferior quality must be interpreted.34 The current study, however, was designed and implemented before the availability of this more recent work. Inevitably, some misclassification will occur, no matter how conservative the cut points are made. Pilot work indicated that by eliminating those patients with displacements in the ambiguous zones, and by requiring two independent raters to classify each patient the same, misclassifications would be minimized. Overall, a patient rated as normal at all three levels was classified as “normal”; patients rated as retro at one or more levels, with no spondy ratings, were classified as retro; and patients rated as spondy at one or more levels were rated as spondy. Additionally information was coded for patients classified as having instability. For those classified as retro, additional information was included that indicated if the retro classification was based on an extension (RE) or a flexion (RF) posture. For those classified as spondy, additional information was included that indicated if the spondy was degenerative (SD) or isthmic (SI).
Treatments. Within each Translation group, patients were assigned to one of three treatment conditions according to a randomization scheme that ensured that, at study completion, each translation classification would be equally represented within each treatment. Each of the three treatment groups included bracing, exercise, and educational elements.

The Flexion Treatment was designed to minimize lumbar extension or lordosis. Each patient was fitted with a Raney Flexion Jacket (Camp International, Inc., Jackson, MI) instructed by a physical therapist in the proper techniques for performing a series of flexion exercises, and viewed an audio/slide presentation emphasizing the value and necessity of avoiding lordotic postures and presenting various coping strategies to avoid lordotic postures.

The Extension Treatment was designed to maintain lumbar extension or lordosis. Each patient was fitted with a Camp hyperextension brace (Camp International, Inc., Jackson, MI) with a pelvic basin, instructed by a physical therapist in the proper techniques for performing a series of Mckenzie-type extension exercises, and viewed an audio/slide presentation emphasizing the value and necessity of maintaining lordotic postures and presenting various coping strategies to maintain lordotic postures.

The Control Treatment was designed to produce no effect and, thus, provided no information regarding flexion or extension postures. Each patient was fitted with an abdominal wrap with Velcro closure. The wrap was designed to hold a thermoplastic molded panel, but no panel was fitted or provided. Each control treatment patient was seen by a physical therapist, but no specific exercise regimen was suggested. If the patient requested an exercise program, the physical therapist advised walking only. These patients viewed an audio/slide presentation summarizing general aspects concerning low-back pain, in which no references to flexion or extension exercise, posturing, or strategies for coping were included.

Compliance. Each flexion and extension brace was modified to include three devices designed to measure compliance. Two timers, one activated by body heat, the other by closure of the brace buckle, and a pedometer were attached to the brace. To encourage compliance, patients were asked to keep a daily log of wearing time and make aware of the three devices on the brace monitoring wear. To help maximize compliance, at 1 week postfitting a nurse contacted each patient by phone to determine how the patient was getting along with wearing the brace and arranged for fitting corrections as necessary.

At 1-month follow-up, a patient was defined as compliant if their self-report log time indicated an effort to wear the brace a minimum of 100 hours if the other compliance measures, the buckle timer, the body temperature timer, the pedometer, had consistent readings (e.g., indicated walking distances consistent with the timer indicators). Technical problems often occurred with timing and pedometer devices and, therefore, as long as at least one of these devices was consistent with the self-reported log entries, and the duration was at least 100 hours, the patient was considered compliant.

Procedure. All new patients received a series of forms to be filled out before arriving at the clinic. When the patient arrived, a clinician evaluated these forms as a means of performing an initial screening of the patient's status relative to demographic and general health study inclusion criteria. If the patient was between the ages of 18 and 60, had back pain for at least 4 weeks but less than 5 years; had not worn any rigid orthosis in the last 6 months; was not pregnant, had no neurologic deficit; was not excessively obese or thin, displayed no gross psychological disturbance; had no respiratory or medical conditions contraindicated bracing; had no history of lumbar fusion; had no pattern of symptoms or physical findings consistent with herniated nucleus pulposus; scoliosis; ankylosing spondylitis; congenital, pathologic or traumatic spondylolisthesis; tumor; vertebral fracture or collapse; discitis; or osteomyelitis, the patient was considered for inclusion in the study. Each patient meeting initial study inclusion criteria was examined using a special physical examination protocol that includes an evaluation for nonorganic signs. If the patient was negative for excessive nonorganic signs, flexion and extension films were evaluated and the patient classified as retro, normal, or spondy. Once within one of the three translation classification categories, the patient was formally asked to participate in the study, informed concerning the nature of the commitment required, and asked to sign an informed consent statement.

The participating patient then was randomly assigned to a treatment. The randomization scheme was designed such that once the study was fully conscripted, patients of each translation type, retro, normal, and spondy, would be assigned in equal numbers to flexion, extension, and control treatments. Once treatment was assigned, the patient was routed to a physical therapist, where range of motion and trunk strengths were assessed and the educational and exercise elements of the treatment were begun.

At the end of this initial visit, the patient was scheduled with an appointment to have the appropriate brace fitted by the orthotist. At the time of the brace fitting, the patient was instructed in the proper use of the brace, and the nature of the various compliance devices was explained. When the fitting was complete, the patient was scheduled for the 1-month follow-up appointment.

At the 1-month follow-up, patients were seen by a clinician who recorded patient's compliance information and a physical therapist evaluated range of motion and trunk strength. In addition, a second set of patient self-report questionnaires were obtained that included outcome measures and patient perceptions of treatment effectiveness.

Design. The study was designed as a $3 \times 3 \times 2$ mixed model repeated measures analysis of variance (ANOVA), with the progress of three translations groups (retro, normal, and spondy) crossed with three treatment modalities (flexion, extension, and control) assessed at admission and at 1-month follow-up. A significant three-way interaction involving these experimental factors (treatment, translation, time) could suggest different patterns of improvement/worsening for the different treatment groups, depending on the nature of the patient's instability, which is, in essence, a restatement of the primary purpose of this study, to demonstrate a distinctive pattern of favorable or unfavorable results, depending on the type of sagittal translation abnormality and treatment modality.

Results

Patient Information
A total of 612 patients were screened for study inclusion over a 2-year period, beginning in August of 1985 and
Table 1. Summary of Selected Personal and Clinical Demographics in Association with Patient’s Translation Classification

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>%</strong></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>26</td>
<td>46%</td>
</tr>
<tr>
<td>Female</td>
<td>30</td>
<td>54%</td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-29</td>
<td>15</td>
<td>27%</td>
</tr>
<tr>
<td>30-39</td>
<td>19</td>
<td>34%</td>
</tr>
<tr>
<td>40-49</td>
<td>13</td>
<td>23%</td>
</tr>
<tr>
<td>50-59</td>
<td>9</td>
<td>16%</td>
</tr>
<tr>
<td>Length of LBP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 6 months</td>
<td>11</td>
<td>20%</td>
</tr>
<tr>
<td>6-12 months</td>
<td>11</td>
<td>20%</td>
</tr>
<tr>
<td>More than a year</td>
<td>34</td>
<td>60%</td>
</tr>
<tr>
<td>Pain from injury?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>25</td>
<td>45%</td>
</tr>
<tr>
<td>Yes</td>
<td>30</td>
<td>55%</td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
<td>5%</td>
</tr>
<tr>
<td>Symptoms started</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gradually</td>
<td>23</td>
<td>41%</td>
</tr>
<tr>
<td>Suddenly</td>
<td>33</td>
<td>59%</td>
</tr>
<tr>
<td>Previous back pain?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>24</td>
<td>44%</td>
</tr>
<tr>
<td>Yes</td>
<td>31</td>
<td>56%</td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
<td>5%</td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working</td>
<td>32</td>
<td>57%</td>
</tr>
<tr>
<td>Not working/LBP</td>
<td>18</td>
<td>32%</td>
</tr>
<tr>
<td>Not working/Other</td>
<td>6</td>
<td>11%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Spondy</th>
<th>Normal</th>
<th>Retro</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>23%</td>
<td>47%</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>0%</td>
<td>16%</td>
<td>47%</td>
</tr>
<tr>
<td></td>
<td>54%</td>
<td>31%</td>
<td>15%</td>
</tr>
<tr>
<td></td>
<td>16%</td>
<td>33%</td>
<td>11%</td>
</tr>
<tr>
<td></td>
<td>55%</td>
<td>36%</td>
<td>38%</td>
</tr>
<tr>
<td></td>
<td>29%</td>
<td>41%</td>
<td>29%</td>
</tr>
<tr>
<td></td>
<td>36%</td>
<td>27%</td>
<td>18%</td>
</tr>
<tr>
<td></td>
<td>32%</td>
<td>32%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>43%</td>
<td>27%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>44%</td>
<td>39%</td>
<td>17%</td>
</tr>
<tr>
<td></td>
<td>27%</td>
<td>36%</td>
<td>36%</td>
</tr>
<tr>
<td></td>
<td>38%</td>
<td>46%</td>
<td>17%</td>
</tr>
<tr>
<td></td>
<td>32%</td>
<td>32%</td>
<td>27%</td>
</tr>
<tr>
<td></td>
<td>34%</td>
<td>31%</td>
<td>34%</td>
</tr>
<tr>
<td></td>
<td>33%</td>
<td>39%</td>
<td>28%</td>
</tr>
<tr>
<td></td>
<td>33%</td>
<td>67%</td>
<td>0%</td>
</tr>
</tbody>
</table>

*Translation Classifications reported as row percentages.

ending in August of 1987. Of these 612 patients, 46% (282/612) were measured for excessive translation in flexion and extension. Overall, 60% (167/282) were classified as “normal,” 24% (67/282) as retro, and 17% (48/282) as spondy. This incidence of excessive translation, approximately 41% for this symptomatic population, suggested that the initial requisite for establishing instability as a clinical syndrome was met. Using the same measurement techniques and classification scheme, Woody et al. reported a significantly lower incidence of instability in an asymptomatic population of men younger than 60 years of age (18%) compared with the 41% incidence of instability in this symptomatic population (P < 0.001). Similarly, the 41% incidence of instability in this symptomatic population was significantly greater than the 15% incidence of instability in a cadaver population reported by Friberg and Hirsch (P < 0.001).

A total of 65 patients met all study inclusion criteria and agreed to participate in the study, and of these 65, 56 returned for a 1-month follow-up evaluation and were, therefore, considered valid participants in the study. A descriptive summary of these 56 patients in terms of gender, age groups, back history, and translation classification is summarized in Table 1.

In general, the sample was relatively evenly divided between men and women, with good representation of patients in their 20s, 30s, 40s, and 50s. Overall, the women in the sample were significantly older than the men (39.9 ± 11 vs. 33.8 ± 8, P < 0.03). Regarding back history criteria, only 20% of the sample had a history of low-back pain of less than 6 months, with 60% reporting low-back pain problems for more than a year; 56% of the sample indicated previous experiences with back pain; and 32% of the sample was not working because of back pain. This pattern suggests that the sample was relatively chronic and, therefore, probably somewhat recalcitrant to treatment. Regarding translation classification, 38.5% (21/56) were classified as normal, 28.5% (16/56) as retro, and 34% (19/56) as spondy. The relative small number of subjects and the lack of equal numbers in each of the translation groups reflects both problems in fully prospecting the study and attrition.

The relationships between translation classification and patient demographics and selected back history criteria, summarized in Table 1, presented an interesting pattern. Translation classification was significantly related to both patient demographic variables, gender and age, but independent with general back pain demographics concerning chronicity, episodic nature of pain, and disability. The pattern of results indicated that women were more likely to be classified as spondy and
men more likely to be classified as retro. Moreover, younger persons were more likely to be classified as retro and older persons as spondy.

**Compliance**

Of the 56 patients who entered the study and returned at 1 month, 18% (10/56) were considered noncompliant. Compliance was significantly related to treatment, with 80% (8/10) of the noncompliant patients in the flexion treatment. Within the flexion treatment, 75% (6/8) of the noncompliant patients were classified as spondy, with the remaining two noncompliant patients classified as normal. Thus, noncompliance was largely restricted to patients classified as spondy in flexion treatment. Despite the disparate nature of noncompliance, several attempts to assess the effects on noncompliance on the various outcome criteria produced no significant effects. Thus, noncompliant patients were not removed from subsequent analyses.

**Treatment Efficacy**

Treatment efficacy was assessed from three perspectives: Patient perceptions of the effectiveness of the brace, exercise, and education programs; effects on trunk strength and range of motion; and overall outcome based on the pain interference (VAS) self-report score.

An initial comparison of patient perceptions of the value of the education, exercise, and bracing components of their treatment program showed no significant differences between treatment groups or with translation categorization. Closer examination of these results, however, suggested that the control treatment group, which received virtually no education or exercise component, responded to these questions inappropriately. After eliminating the control treatment from the analysis, significant differences in perceived benefits were observed between the extension and flexion treatments ($F_{2,60} = 3.65, P < 0.04$). Examination of the pattern of perceived benefits accorded to the education, exercise, and brace components of each treatment indicated that the flexion treatment group reported relatively low benefits from all treatment components whereas the extension treatment group reported relatively large benefits from the brace and the education components of the treatment.

Comparison of patient total range of motion and trunk strength from the study admission to 1-month follow-up indicated that neither total range of motion or trunk strength was adversely affected by treatment. Contrary to expectation, significant increases, rather than decreases, in total range of motion and flexion trunk strength were observed, $F_{1,44} = 4.22, P < 0.03$, and $F_{1,44} = 8.51, P < 0.01$, respectively. From admission to 1-month follow-up, average total range of motion increased from 62.4° to 66.0°, and flexion trunk strength increased from 34.9 to 40.1 kg. Although there were no significant interactions suggesting differential improvement across treatments or translation categories, there was a trend suggesting increased flexion trunk strength was more likely in both the flexion and extension treatment groups compared with the control group, with relative gains of 6.59 and 7.52 kg compared with 0.83 kg, $F_{2,44} = 3.03, P < 0.06$.

The primary measure of treatment efficacy was the 15-item pain interference (VAS) scale. Overall there was significant improvement across time, $F_{1,44} = 6.80, P < 0.02$. At admission, the average VAS score was 5.5 (range, 2.3–8.3) and had a coefficient alpha internal consistency reliability estimate of 0.835. At 1-month follow-up, the average VAS score was 6.0 (range, 3.3–9.9), and had a reliability of 0.884. A significant interaction between treatment and time, $F_{1,47} = 4.03, P < 0.03$, however, indicated that improvement was differential across treatments. The pattern of VAS scores across time by treatment groups is shown in Figure 1. Follow-up tests of simple effects indicated no significant differences between treatments at admission, $F_{2,47} = 1.62, P < 0.21$, but significant treatment differences at 1 month, $F_{2,47} = 4.14, P < 0.03$. At 1-month follow-up, the extension VAS score of 6.85 was significantly greater than the flexion and control treatment groups’ mean VAS scores of 5.48 and 5.97, respectively. In addition, only the extension treatment showed significant improvement across time (improvement from 5.6 to 6.85, $F_{1,15} = 11.61, P < 0.004$). Thus, the overall improvement across time was predominantly reflective to improvement in extension treatment.

**Discussion**

**Instability as a Clinical Syndrome**

The primary purpose of this study was to determine if flexion or extension treatments would result in a distinctive pattern of favorable or unfavorable results depending on the type of sagittal translation abnormality. Even when considering the problems associated with
assessing radiographic instability and the limitations in this study, these data provided little support for the hypothesis. The greater incidence of radiographic instability in the symptomatic populations (41%) compared with an asymptomatic sample (18%) was the only aspect of the results supportive of the notion that instability may represent a clinical syndrome.

Given the small numbers in each subgroup resulting from jointly considering translation classification and treatment modality and the resultant low power of the statistical tests evaluating the translation by treatment effects across time, it was decided to graph treatment effects across time for each of five translation subgroups (i.e., normal, excessive retrodisplacement in flexion [RF] and extension [RE], degenerative spondylolisthesis [SD], and isthmic spondylolisthesis [SI]). The pattern of treatment effects (control, flexion, and extension) across time were virtually identical for all five translation subgroups. Thus, the lack of differential treatment effects depending on translation classification was not considered a consequence of small sample size and a resultant lack of power for identifying truly different treatment effects across translation categories over time.

The failure of the flexion and extension treatments to produce a pattern of outcome useful for corroborating radiographic findings of excessive translation in the sagittal plane was disappointing, but, in retrospect, not surprising. Given the long history of “instability” as a recognized condition with a clearly indicated treatment, lumbar fusion, the lack of any forthcoming corroborating sign should suggest caution. Although a number of clinical signs for corroborating instability have been suggested, there has been very little systematic and scientific study of the relationship between these signs and radiographic instability. Stokes and Frymoyer attempted to establish relationships between clinical signs and radiographic instability without success. The inability to isolate and demonstrate specific clinical signs that correlate with radiographic instability suggests that the clinical syndrome of instability may be illusory. There may be no corroborative signs.

At least three problems have combined to make the concept of “instability” difficult to isolate as a clinical syndrome. Perhaps most fundamentally, attempts to carefully quantify instability radiographically have ignored or forgotten the inherent aspects of the concept. The original description of abnormal translation on flexion extension radiographs by Knutsson and the patho-anatomic correlations with these radiographic findings described by Friberg and Hirsch emphasize that the findings were associated with disc pathology. It may not make sense to separate instability from the evaluation of the disc. In considering the model of instability proposed by Kirkaldy-Willis and Farfan, excessive translation is expected most commonly in the “late phase” of instability. That is, in chronic low-back pain populations, abnormal alignment or deformity is most commonly detected on flexion–extension radiographs. Motion segments with these deformities usually do not exhibit excessive excursion. If the “early phase” of instability is characterized by early abnormalities in the annulus and nucleus of the disc, it may not be possible to detect “early phase” instability from flexion–extension radiographs. It is not unlikely that the three translation groups in this study had a common underlying disc pathology. The type of translation abnormality by which patients were classified may be secondary in determining treatment response. Because all three translation groups (spondylolisthesis, normal, and retro) may have had the similar underlying pathologies, it would not be unexpected that they may have similar responses to treatment. Specifically, all three groups responded favorably to extension or lordotic treatment, which is a treatment proposed for patients with discogenic low-back pain. Thus, radiographic instability, as described in this study, may need to be considered as a “sign” of disc degeneration rather than a distinct clinical syndrome. Or, if considered a clinical syndrome, instability would need to be evaluated and corroborated by evidence of disc pathology such as decreased hydration on T2-weighted magnetic resonance imaging (MRI) images or positive discography. Quinnell and Stockdale demonstrated a high correlation between abnormal anteroposterior translation and abnormal discographic appearance. Because clinical symptoms and signs to corroborate instability are lacking, it may be rational to reconsider the value of the pain response during discography as a corroborating sign for instability. Although the clinical value of discography has been debated, recent evidence has demonstrated a low false-positive rate for discography in normal subjects when the pain response is considered.

A second problem concerns differences in how translation measured are evaluated to diagnose or classify “instability.” Two basic approaches have been established: The maximum translation approach, used in this study, in which the largest amount of translation in flexion or extension is used to classify the patient; and the excursion approach, in which the difference in translation in flexion and extension postures is used to classify the patient. The possibility that these two approaches represent different constructs or that these methods serve to differentiate between patients at different stages of the same process, where patients displaying large excursions may reflect earlier stages of instability and patients displaying maximal translations with little excursion may reflect later stages of instability, suggests that jointly considering both definitions may provide a more useful classification of instability. For the excursion approach to prove useful, however, traditional standing flexion–extension roentgenograms may need to be abandoned in favor of newer methods, because flexion–extension roentgenograms in the upright position do not generally exhibit excursion or hypermobility in the sagittal plane.
Recent methods studying the effect of traction and compression on the patient while obtaining lateral radiographs have demonstrated significant hypermobility in some patients with spondylolisthesis.\textsuperscript{8,15} Alternatively, more accurate assessment of intervertebral motion by the use of surgically implanted markers, as described by Selvik,\textsuperscript{33} may provide more valid measurement of translation. Perhaps “instability” or hypermobility as defined by one of these methods will yield better correlation with symptoms, clinical signs, and treatment response. Conceptually, the presence of hypermobility may be clinically important to the treating physician as a means of assessing instability or disc degeneration in its earlier phases. The notion that the excursion and maximum translation approaches to radiographic instability might be differentially sensitive to the progression of discogenic problems suggests that longitudinal studies incorporating MRI techniques might be useful for documenting the natural history of discogenic problems. If a relationship between radiographic instability and the progression of discogenic problems can be established, then the door is opened for systematic investigation of the relative efficacy of early versus late interventions. For example, if hypermobility is an early sign of discogenic problems, then perhaps spinal fusion can be averted by implementing a comprehensive conservative treatment regimen.

A third problem concerns the dependence of radiographic instability on the evaluation of clinical roentgenograms and the inherent problems in obtaining reliable and valid sagittal translation measurements.\textsuperscript{34} Three particular issues have been raised. When film quality is less than optimal and concomitant motions exist in the other planes of the spine, accurate measures of translation are difficult to obtain. A second issue concerns differences in the methods used to measure translation. Of the several methods in common use for measuring sagittal translation, some are clearly more accurate than others. A third issue concerns a lack of clear and defensible standards specifying how much translation is normal and how much is excessive.

**Treatment Efficacy**

The significant treatment by time interaction, and the stability of this interaction over all translations was somewhat surprising. The form of this interaction indicated general improvement at 1 month over admission for patients in the extension treatment, compared with virtually no improvement for the control and flexion treatment groups. Unfortunately, the nature of the treatment groups, which combined bracing, exercise, and education components in an attempt to maximize treatment effect, simultaneously confounds the ability to determine what effects each of these components may have had on the outcome. Was the significant benefit of the extension treatment due to the brace, the exercise, the education, or some combination of these three and, if so, what contribution from each? Patients’ perceptions of the relative benefits for these three components of treatment suggested that exercise was not considered of great benefit in either the flexion or extension treatments. Patients in extension treatment, however, did report greater perceived benefits from the education and bracing treatment components, compared with the patients treated in flexion. Unfortunately, perceptions of benefit and actual efficacy need not be causally linked.

Positive results from extension treatment for patients within “normal” translation limits was not unexpected; others have reported such favorable results.\textsuperscript{36,41} Positive results from extension treatment for patients with spondylolisthesis was unexpected, however, and, in fact, it was initially hypothesized that this treatment should actually exacerbate symptoms for this population. A previous report of an athlete population reported that patients with spondylolysis or spondylolisthesis responded favorably when treated with a rigid flexion orthosis.\textsuperscript{23} In addition, because both degenerative and isthmic types of spondylolisthesis are associated with posterior element disease, one might expect a relative lateral recess or foraminal stenosis in the extension posture for these patients. That retrodisplacement patients responded favorably to extension treatment was also surprising because, in most situations, the posterior translation was maximized on the extension roentgenogram. Therefore, extension treatment should have increased the deformity, thereby exacerbating the symptoms.

One explanation of the nondifferential effects of extension treatment across types of radiographic instability is that, for this population, as opposed to a population of young male athletes, radiographic instability was secondary to an underlying disc abnormality or pathology. Thus, although the patients were classified according to the nature of their radiographic instability, they were actually more common than different when considering the underlying or primary cause or causes for their instability. Therefore, a treatment designed to affect the dynamics (muscle forces, disc nutrition, intradiscal pressure) of the functional spinal unit was beneficial, regardless of the consequence of the disc abnormality. A second explanation for improvement with extension treatment for the retrodisplacement and degenerative spondylolisthesis patients might be that extension posture “locks” the facet joints and thus stabilizes the “unstable” symptomatic motion segments. This latter hypothesis, however, provides no explanation for why the isthmic spondylolisthesis patients would show improvement in the extension treatment because the discontinuity in the posterior elements would not allow “locking” and the resultant stability to occur. In general, it would seem that the clinician considering a rigid orthosis as
treatment for chronic low-back pain should consider extension bracing, in conjunction with an extension exercise program and posture program, regardless of whether radiographic signs suggest spondylolisthesis or retrodisplacement.

**Limitations**

A number of factors combined to limit the generalizability of this study. First, the inclusion and exclusion criteria established for this study proved to be much more restrictive than expected. Thus, of 612 patients screened, only 65 met all inclusion criteria and were willing to participate, and of these, only 56 returned at 1 month so that treatment effects could be assessed. Small sample size limits the generalizability of any results because smaller groups tend to be less representative of target populations, random assignment to conditions can be less successful, and the statistical power of the analytical procedures is limited.

A lower than expected incidence of suitable subjects was noted early in the study. Steps were taken to increase the sample size by liberalizing some inclusion criteria, specifically by allowing patients with a duration of pain greater than 1 year to enter the study. The result was that the sample took on a more chronic low-back trouble character. For example, 30% of the sample (17/56) reported being disabled (unable to work) because of back pain. Thus, the recalcitrant nature of this sample undoubtedly made it more difficult to establish the efficacy of the bracing treatments. 39

Another limitation concerns the variation in compliance. Patients who wore their brace less often may not have received the full relief or exacerbation of symptoms that might have been possible with greater compliance, therefore limiting the ability of true treatment effects to be expressed. The data were analyzed to assess the effects of compliance, however, and no significant differences in patterns of results were observed when comparing results including and excluding noncompliant patients. In addition, the finding that noncompliance was more likely in the flexion treatment relative to the extension treatment and for patients classified as spondy relative to those classified as retro or normal remains a concern even, though patterns of progress across time were not significantly different when noncompliant subjects were removed from analyses.

Finally, the 1-month follow-up period may have been of insufficient length to establish any practical treatment effects. In defense of this short follow-up period, however, the primary purpose of the bracing treatments was to evaluate their value as corroborating diagnostic tests for radiographic instability. Because it was hypothesized that some randomly assigned treatments for the various instability types might actually exacerbate symptoms, and the worsening effects were considered as important as beneficial effects in terms of establishing corroborative signs, it was determined that 1 month was a sufficient follow-up. Because it was expected that some braces would be relatively uncomfortable and ineffective, noncompliance and attrition were concerns and might lead to loss of important information if follow-up was set at a longer interval. Furthermore, several clinicians have indicated that bracing treatments should be of relatively short duration. 3, 6, 11, 26

**Summary**

Although this study provided little evidence to suggest that radiographic instability represents a viable clinical syndrome, evidence was presented suggesting that extension treatment, including bracing, education, and exercise components, was effective. Significant reductions in pain interference at 1-month follow-up were demonstrated. This treatment effect should be considered all the more significant given the arguments set forth regarding the recalcitrant nature of this sample, and the questionable strength of the bracing treatments (i.e., the short duration of treatment, incidence of noncompliance, and the minimal requirements necessary to be considered compliant). The practicing clinician might consider short-term extension bracing as a viable conservative treatment for patient presenting with low-back pain, regardless of whether the patient has chronic pain or radiographic instability.

This study served to illustrate a wide range of problems inherent in investigating radiographic instability. The restrictive definition of radiographic instability, which focuses only on translation in the sagittal plane, and the disagreement among clinicians concerning whether radiographic translation should be measured according to excursion or maximum translation, indicates that, although the concept of radiographic instability appears well established, it has not been adequately operationalized. In general, the lack of precision associated with radiographic instability speaks to the subjective nature of the diagnosis and suggests that the clinician must exercise caution when basing treatment decisions on this radiographic abnormality.

This was the first prospective, randomized study attempting to establish radiographic instability as a clinical syndrome. The limitations of the study as well as fundamental problems in the general methodology in this area have been summarized in the hope that future investigators will benefit. The implications of the effectiveness of extension bracing treatments are apparent. Given the relatively recalcitrant nature of this sample, the effectiveness of the extension brace in reducing pain interference suggests that extension bracing, with complimentary education and exercise programs, may represent a relatively powerful modality that might be considered a viable conservative treatment approach, even for patients with chronic pain.
References


Address reprint requests to
Kevin F. Spratt, PhD
Spine Diagnostic & Treatment Center
1075 Roy Carver Pavilion
Department of Orthopaedic Surgery
University of Iowa Hospitals & Clinics
Iowa City, IA 52242