

A Clinical Prediction Rule To Identify Patients with Low Back Pain Most Likely To Benefit from Spinal Manipulation: A Validation Study

Maj John D. Childs, PhD, PT; Julie M. Fritz, PhD, PT; Timothy W. Flynn, PhD, PT; James J. Irrgang, PhD, PT; Maj Kevin K. Johnson, PT; Maj Guy R. Majkowski, PT; and Anthony Delitto, PhD, PT

Background: Conflicting evidence exists about the effectiveness of spinal manipulation.

Objective: To validate a manipulation clinical prediction rule.

Design: Multicenter randomized, controlled trial.

Setting: Physical therapy clinics.

Patients: 131 consecutive patients with low back pain, 18 to 60 years of age, who were referred to physical therapy.

Intervention: Patients were randomly assigned to receive manipulation plus exercise or exercise alone by a physical therapist for 4 weeks.

Measurements: Patients were examined according to the clinical prediction rule criteria (symptom duration, symptom location, fear-avoidance beliefs, lumbar mobility, and hip rotation range of motion). Disability and pain at 1 and 4 weeks and 6 months were assessed.

Results: Outcome from spinal manipulation depends on a patient's status on the prediction rule. Treatment effects are greatest for the subgroup of patients who were positive on the rule (at least 4 of 5 criteria met); health care utilization among this subgroup was decreased at 6 months. Compared with patients who

were negative on the rule and received exercise, the odds of a successful outcome among patients who were positive on the rule and received manipulation were 60.8 (95% CI, 5.2 to 704.7). The odds were 2.4 (CI, 0.83 to 6.9) among patients who were negative on the rule and received manipulation and 1.0 (CI, 0.28 to 3.6) among patients who were positive on the rule and received exercise. A patient who was positive on the rule and received manipulation has a 92% chance of a successful outcome, with an associated number needed to treat for benefit at 4 weeks of 1.9 (CI, 1.4 to 3.5).

Limitations: The response rate for the 6-month follow-up resulted in inadequate power to detect statistically significant differences for some comparisons.

Conclusions: The spinal manipulation clinical prediction rule can be used to improve decision making for patients with low back pain.

Ann Intern Med. 2004;141:920-928.

www.annals.org

For author affiliations, see end of text.

The trial will be registered in Current Controlled Trials (www.controlled-trials.com) in the near future.

See related articles on pp 901-910 and pp 911-919 and editorial comment on pp 957-958.

Next to the common cold, low back pain is the most common reason that individuals visit a physician's office (1). Billions of dollars in medical expenditures and lost labor costs for this condition are incurred each year (2, 3). Attempts to identify effective interventions for individuals with low back pain have been largely unsuccessful (4). In particular, conflicting evidence exists about the effectiveness of spinal manipulation; some randomized trials have shown a benefit, while other trials have not (5-7). These conflicting conclusions are reflected in the various recommendations in national clinical practice guidelines, with some recommending manipulation and others not (8).

The variety of conclusions in trials of manipulation may be attributable to the failure of researchers to adequately consider the importance of classification. Using broad inclusion criteria results in a heterogeneous sample that may include many patients for whom no benefit is expected, thus masking the intervention's true value (9, 10). Consequently, developing methods for matching patients with low back pain to treatments that are most likely to benefit them has become an important research priority (11). Clinical prediction rules are tools designed to assist clinicians in decision making when caring for patients (12). Several clinical prediction rules have been developed and validated to improve clinical decision making

for the use of imaging in patients with ankle, knee, cervical spine, or minor head injuries (13-16). Few studies have attempted to develop rules that establish prognosis on the basis of outcome from a specific intervention, such as spinal manipulation.

Recently, Flynn and colleagues (17) developed a clinical prediction rule for identifying patients with low back pain who are likely to benefit from manipulation. They examined a series of patients with low back pain who received a manipulation intervention. Five factors formed the most parsimonious set of predictors for identifying patients who achieved at least 50% improvement in disability within 1 week with a maximum of 2 manipulation interventions (Table 1) (17). The positive likelihood ratio among patients who met at least 4 of 5 of the criteria was 24.4 (95% CI, 4.6 to 139.4).

Clinical prediction rules must be validated in separate populations before being recommended for widespread implementation (18). A clinical prediction rule for identifying which patients with low back pain are most likely to respond to manipulation could improve clinical efficiency and resource utilization. Thus, we aimed to validate the spinal manipulation clinical prediction rule in a multicenter trial.

METHODS

We considered consecutive patients with a primary symptom of low back pain who were referred to physical therapy for participation. We used 14 physical therapists at 8 clinics in various U.S. regions and settings (2 academic medical centers and smaller outpatient practice settings). Most participating sites were health care facilities within the U.S. Air Force. Each site's institutional review board approved the study before we began recruitment and data collection.

Inclusion criteria were age 18 to 60 years; a primary symptom of low back pain, with or without referral into the lower extremity; and an Oswestry Disability Questionnaire (ODQ) score of at least 30%. We excluded patients who had "red flags" for a serious spinal condition (for example, tumor, compression fracture, or infection), those who had signs consistent with nerve root compression (that is, positive straight-leg increase < 45 degrees or diminished reflexes, sensation, or lower-extremity strength), those who were pregnant, or those who had previous surgery to the lumbar spine or buttock. These criteria are consistent with those used in Flynn and colleagues' study (17) and were designed to include patients without a contraindication to manipulation. Once patients were admitted to the study, we used intention-to-treat principles, and no patient was removed for nonadherence.

History and Physical Examination

Before randomization, patients completed several self-report measures and then received a standardized history and physical examination. We collected demographic information, including age and sex; medical history; and location and nature of symptoms. Self-report measures included a body diagram to assess the symptom distribution (19). We used an 11-point pain-rating scale ranging from 0 (no pain) to 10 (worst imaginable pain) to assess current pain intensity and the best and worst level of pain during the last 24 hours (20). We used the average of the 3 ratings. We used the Fear-Avoidance Beliefs Questionnaire (FABQ) to quantify the patient's fear of pain and beliefs about avoiding activity (21). Previous studies have found a high level of test-retest reliability for both the FABQ phys-

Table 1. Five Criteria in the Spinal Manipulation Clinical Prediction Rule*

Criterion	Definition of Positive
Duration of current episode of low back pain	<16 d
Extent of distal symptoms	Not having symptoms distal to the knee
FABQ work subscale score	<19 points
Segmental mobility testing	≥1 hypomobile segment in the lumbar spine
Hip internal rotation range of motion	≥1 hip with >35 degrees of internal rotation range of motion

* See Appendix 1 and Appendix 3 video, available at www.annals.org, for details (17). FABQ = Fear-Avoidance Beliefs Questionnaire.

Context

In this randomized, controlled trial, spinal manipulation plus exercise produced outcomes for low back pain similar to those produced by exercise alone. Yet, some patients did respond to spinal manipulation, and it would be helpful for doctors to be able to identify such patients.

Contribution

Patients were most likely to benefit from spinal manipulation if they met 4 of 5 of the following criteria: symptom duration less than 16 days, no symptoms distal to knee, score less than 19 on a fear-avoidance measure, at least 1 hypomobile lumbar segment, and at least 1 hip with more than 35 degrees of internal rotation.

Implications

Clinicians may be able to use these criteria to identify patients with low back pain who are good candidates for spinal manipulation.

—The Editors

ical activity and work subscales (22). Fear-avoidance beliefs have been associated with current and future disability and work loss in patients with acute (23) and chronic (24) low back pain. The modified ODQ is a region-specific disability scale for patients with low back pain (25) that has high levels of reliability, validity, and responsiveness (26).

Physical examination measures included lumbar active range of motion (27) and various tests purported to identify dysfunction in the lumbopelvic region (28). Complete details of the physical examination are described elsewhere (26). Specific components pertinent to validation of the rule were assessments of segmental mobility and hip internal rotation range of motion, the performance of which is described in Appendix 1 and Appendix 3 video (available at www.annals.org). Each physical therapist received a detailed manual that operationally defined each examination and treatment procedure and was trained in the study procedures by an investigator before data collection began.

Determining Status on the Clinical Prediction Rule

A physical therapist who was blinded to the patients' treatment group assignment assessed the 5 criteria in the rule (Table 1, Appendix 1, and Appendix 3 video). To further minimize bias, examiners were not instructed in the rule's criteria and were unaware of the patient's status on the rule. After completion of the study, an examiner who was blinded to the patient's treatment assignment determined the patient's status on the rule by using the results of the baseline examination. As was done in the initial study (17), we classified patients as positive if they met at least 4 of 5 criteria and were therefore likely to respond to manipulation. We classified patients with 3 or fewer criteria as negative. An examiner who was blinded to the patient's status on the rule repeated the history and physical

Figure 1. Manipulative intervention used in developing and validating the spinal manipulation clinical prediction rule.



See Appendix 2 and Appendix 3 video, available at www.annals.org, for details. Reprinted from reference 30: Childs JD, Fritz JM, Piva SR, Erhard RE. Clinical decision making in the identification of patients likely to benefit from spinal manipulation: a traditional versus an evidence-based approach. *J Orthop Sports Phys Ther.* 2003;33:259-75, with permission of the Orthopaedic and Sports Physical Therapy Sections of the American Physical Therapy Association.

examination 1 and 4 weeks after randomization. Patients also completed a 6-month follow-up postal questionnaire to assess disability, work status, and health care utilization.

Treatment Groups

We used a random-number generator to generate a randomization list before the study began. We prepared individual, sequentially numbered index cards with the randomization assignments. We folded the cards and placed them in sealed envelopes. After the baseline examination, the physical therapist who conducted the examination opened the next envelope, indicating the treatment group assignment. We randomly assigned patients to 1 of 2 groups: 1) spinal manipulation plus an exercise program (manipulation group) or 2) an exercise program alone (exercise group). Patients in both groups attended physical therapy twice during the first week and then once a week for the next 3 weeks, for a total of 5 sessions. We initiated treatment immediately after completion of the baseline examination, unless prohibited by time constraints; in that case the first treatment session took place 24 to 48 hours after the baseline examination. All patients received an exercise instruction booklet that outlined the proper performance and frequency of each exercise and were instructed to perform their assigned exercise program once daily on the days

that they did not attend therapy. On the basis of the benefits associated with remaining active (29), patients in both groups were given advice to maintain usual activity within the limits of pain.

Manipulation Group

The treatment received by the manipulation group differed from that of the exercise group during the first 2 physical therapy sessions. During these 2 sessions, patients received high-velocity thrust spinal manipulation and a range-of-motion exercise only. First, the physical therapist performed the manipulation by using the same technique used by Flynn and colleagues (17). Appendix 2 (available at www.annals.org) describes and Figure 1 and Appendix 3 video illustrate the procedures used to perform the manipulation technique.

Exercise Group

We treated patients in the exercise group with a low-stress aerobic and lumbar spine strengthening program. The Agency for Health Care Policy and Research (AHCPR) clinical practice guidelines for Acute Lower Back Problems in Adults (31) recommends muscle-strengthening exercises

for patients with acute low back pain, and evidence also supports exercise therapy for patients with chronic low back pain (32). The strengthening program was designed to target the trunk musculature identified as important stabilizers of the spine in the biomechanical literature (33). The AHCPR guidelines (31) also recommend low-stress aerobic exercises for patients with acute low back pain; thus, we also included an aerobic exercise component. Patients began with a goal of 10 minutes of aerobic exercise on a stationary bike or treadmill at a self-selected pace. The exercise program progressed according to criteria previously described (34).

Sample Size Determination

We based sample size calculation on detecting a statistically significant 3-way interaction between a patient's status on the rule, treatment group, and time by using the 1-week ODQ score at an α level of 0.05. The study was powered on the interaction because its detection would contribute most significantly to the validity of the rule. On the basis of previous research (35, 36), we expected a within-cell SD of 15 points on the ODQ score and a correlation between the covariate and dependent variable of 0.30 ($R^2 = 0.09$). Given these variables, we required 21 patients per cell for a moderate effect size (0.30) for the interaction with 80% power by using a 2-tailed hypothesis. In the previous study (17), 30% of patients were positive on the rule. Assuming a similar distribution, we required approximately 130 to 140 patients to assure that 21 patients who were positive on the rule would be randomly assigned to each group.

Statistical Analysis

We compared baseline variables between groups by using independent *t*-tests or Mann-Whitney U tests for continuous data and chi-square tests of independence for categorical data. We examined the primary aim with a 3-way repeated-measures multivariate analysis of variance (ANOVA) with treatment group (manipulation vs. exercise) and status on the clinical prediction rule (positive or negative) as between-patient variables and time (baseline, 1 week, 4 weeks, and 6 months) as the within-patient variable. The dependent variables were disability (ODQ score) and pain. The hypothesis of interest was the 3-way interaction. The primary outcome measure was the 1-week ODQ score to mirror the follow-up used in Flynn and colleagues' study (17). We assessed the ODQ score after 4 weeks and again after 6 months to determine whether a patient's status on the rule predicted outcome at a longer follow-up. We assessed self-reported levels of pain to determine whether similar changes in pain occurred. We performed planned pairwise comparisons at each follow-up period by using the Bonferroni inequality (37). We hypothesized that patients who were positive on the rule and received manipulation would experience greater improvement in 1- and 4-week outcomes than patients who were negative on the rule and received manipulation, compared with patients who were positive on the rule but received

the exercise program without manipulation. Because the clinical prediction rule is believed to be specific to a manipulation intervention, we hypothesized that outcomes among those receiving the exercise treatment would not differ on the basis of a patient's status on the rule. Given the favorable natural history of low back pain, we hypothesized that outcomes would not differ on the basis of a patient's status on the rule after 6 months.

To elucidate the value of incorporating the rule into clinical practice, each patient's outcome was dichotomized as successful or unsuccessful on the basis of the percentage change in the ODQ scores at the 1- and 4-week follow-ups. We classified patients with at least 50% improvement as successful and all others as unsuccessful. We constructed hierarchical logistic regression models to examine the relationship between treatment group and status on the prediction rule with the dichotomized outcomes after controlling for other variables. We performed regression modeling in 3 steps. We constructed separate models for the 1-week and 4-week outcomes. We entered potentially confounding baseline factors in the first step (age, sex, body mass index, history of low back pain, previous response to manipulation, and baseline disability score). Second, we entered the treating therapist to control for potential differences related to the clinician. We entered treatment group, prediction rule status, and the interaction between treatment and rule status in a stepwise manner in step 3. A significance level of *P* less than 0.05 was required for entry.

For patients randomly assigned to the manipulation group, we calculated sensitivity, specificity, and positive and negative likelihood ratios to describe the accuracy of the rule's criteria for predicting treatment success after 1 week with spinal manipulation (38). We also calculated number needed to treat statistics to further illustrate the value of the rule to clinicians. We used intention-to-treat principles to account for participants who dropped out.

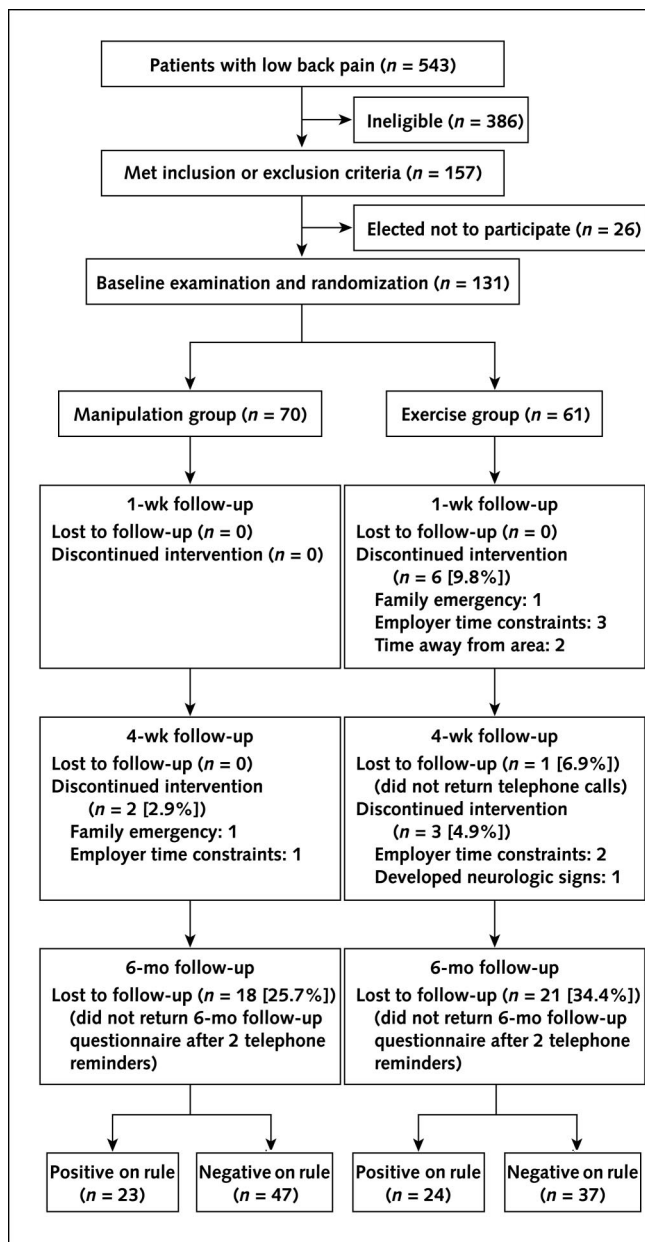
Role of the Funding Sources

The Foundation for Physical Therapy, Inc., and the Wilford Hall Medical Center Commander's Intramural Research Funding Program supported the study. These agencies played no role in the design, conduct, or reporting of the study or in the decision to submit the manuscript for publication.

RESULTS

We recruited patients from 8 clinics in various U.S. regions from March 2002 through March 2003 (Figure 2). We screened 543 consecutive patients who were referred to physical therapy with symptoms of low back pain for inclusion. Of these, 386 did not meet all inclusion criteria. The most common reasons for exclusion were an ODQ score less than 30% ($n = 202$ [53%]) and age younger than 18 years or older than 60 years ($n = 64$ [17%]). Of 157 eligible patients, 26 elected not to participate (Figure 2) because of concerns about the time commitment ($n = 19$) or not wanting to be randomly assigned to 1 of

Figure 2. Flow diagram for patient recruitment and randomization.



the treatment groups ($n = 7$). The remaining 131 patients provided informed consent and were enrolled. We randomly assigned 70 patients to the manipulation group and 61 to the exercise group (Figure 2).

According to intention-to-treat principles, we included all 131 patients in the analysis by carrying forward the last observation. Baseline variables did not differ between treatment groups (Table 2). More patients dropped out of the exercise group before both the 1-week (6 vs. 0 patients) and 4-week (4 vs. 2 patients) follow-ups ($P = 0.007$). All patients reported non-study-related reasons for dropping out (such as time constraints and family issues). The overall

response rate for the 6-month follow-up was 70.2%. The response rates between patients in the manipulation (74.3%) and exercise (65.6%) groups did not differ ($P > 0.2$).

The overall 3-way clinical prediction rule \times treatment group \times time interaction for the repeated-measures multivariate ANOVA was statistically significant, indicating that the outcome depended on both the patient's treatment group and status on the rule. We observed a similar pattern for the pain scores. Post hoc comparisons demonstrated that patients who were positive on the rule and received manipulation experienced greater improvement in 1- and 4-week disability than patients who were negative on the rule and received manipulation. This difference was maintained at the 6-month follow-up (Table 3, Figure 3). Furthermore, patients who were positive on the rule and received manipulation also experienced greater improvement in 1- and 4-week disability than patients who were positive on the rule but received the exercise intervention. This difference was also maintained at the 6-month follow-up (Table 3, Figure 3). Among patients in the exercise group, we observed no differences in the patient's status on the rule at the 1-week, 4-week, or 6-month follow-ups (Table 3, Figure 3). Post hoc comparisons of pain scores demonstrated a similar pattern. Regardless of a patient's status on the rule, patients who received manipulation experienced greater improvements in disability and pain than those who did not; however, except for the 6-month follow-up, treatment effects were generally smaller than when a patient's status on the rule is considered (Table 3).

At the 6-month follow-up, patients in the exercise group demonstrated statistically significantly greater medication use, health care utilization, and lost work time due to back pain than patients in the manipulation group. Among patients who were positive on the rule, a greater proportion of patients in the exercise group were currently seeking other treatment for their back pain (Table 4).

After 1 week, 31 of 70 (44.3%) patients in the manipulation group achieved success compared with 7 of 61 (11.5%) patients in the exercise group ($P < 0.001$). After 4 weeks, 44 of 70 (62.9%) and 22 of 61 (36.1%) patients achieved success in the manipulation and exercise groups, respectively ($P = 0.002$). For the hierarchical logistic regression analysis in which the 4-week outcome was the dependent variable, the entry of baseline factors ($P > 0.2$) and treating therapist ($P = 0.12$) did not significantly contribute to the prediction of success. The regression analysis only entered the interaction between treatment group and prediction rule status in the third step ($P < 0.001$). Relative to the reference category of patients who were negative on the rule and received exercise, the adjusted odds ratio among patients who were positive on the rule and received manipulation was 60.8 (CI, 5.2 to 704.7; $P = 0.001$). The adjusted odds ratios among patients who were negative on the rule and received manipulation (2.4 [CI, 0.83 to 6.9]) and positive on the rule and received exercise (1.0 [CI, 0.28 to 3.6]) did not reach significance in the final model.

Table 2. Baseline Demographic and Self-Reported Variables for Both Treatment Groups*

Variable	All Patients (n = 131)	Manipulation Group (n = 70)	Exercise Group (n = 61)	P Value
Age, y	33.9 ± 10.9	33.3 ± 11.2	34.6 ± 10.6	>0.2
Women, %	42	42.9	41.0	>0.2
Body mass index, kg/m ²	27.1 ± 4.5	27.7 ± 4.7	26.3 ± 4.1	0.08
Current smokers, %	22.9	17.1	29.5	0.09
History of low back pain, %	67.9	65.7	70.5	>0.2
Previous improvement with manipulation for low back pain, %	22.1	15.7	29.5	0.06
Median duration of current symptoms, d	27	22	30	>0.2
Medication use for low back pain, %	84.0	87.1	80.3	>0.2
Missed any work for low back pain, %	39.8	41.2	38.3	>0.2
Symptoms distal to the knee, %	23.7	25.7	21.3	>0.2
FABQ physical activity subscale score	17.0 ± 4.3	16.8 ± 4.9	17.3 ± 3.5	>0.2
FABQ work subscale score	17.0 ± 10.3	16.5 ± 10.1	17.4 ± 10.5	>0.2
ODQ score	41.2 ± 10.4	41.4 ± 10.1	40.9 ± 10.8	>0.2
Pain score	5.8 ± 1.6	5.7 ± 1.7	5.9 ± 1.5	>0.2

* Values presented with a plus/minus sign are means ± SD. FABQ = Fear-Avoidance Beliefs Questionnaire; ODQ = Oswestry Disability Questionnaire.

We observed the same pattern for 1-week follow-up, with an adjusted odds ratio of 114.7 (CI, 11.4 to 1155.0) ($P < 0.001$) among patients who were positive on the rule and received manipulation.

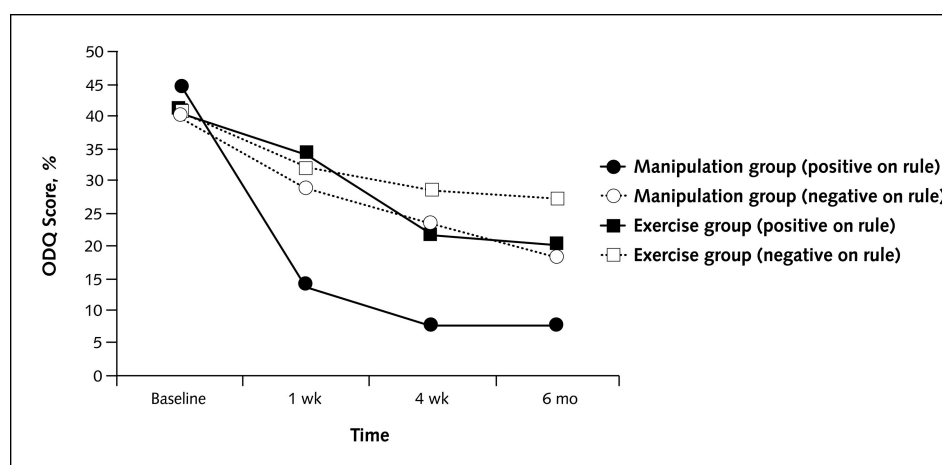
In the manipulation group, positive status on the rule resulted in a positive likelihood ratio of 13.2 (CI, 3.4 to 52.1) for predicting success at 1 week. The negative likelihood ratio associated with meeting fewer than 3 criteria was 0.10 (CI, 0.03 to 0.41). Among patients who were positive on the rule, the number needed to treat for benefit for a successful outcome was 1.3 (CI, 1.1 to 1.9) at 1 week with manipulation and 1.9 (CI, 1.4 to 3.5) at 4 weeks with manipulation.

DISCUSSION

Current evidence supports only a few common interventions for patients with low back pain, including non-steroidal anti-inflammatory medications (39), advice to remain active (40), and exercise (41). Conflicting results

from randomized trials and systematic reviews on the effectiveness of manipulation suggest that while some patients respond rather dramatically, others may not experience much improvement. Previous studies that have not attempted to identify the subgroup of patients who are likely to benefit from manipulation have generally shown small treatment effects, leading to questions about the effectiveness of the intervention (6). Our results confirm that patients with low back pain who are likely to benefit from manipulation can be identified with increasing certainty. Patients who were positive on the rule and treated with manipulation had greater improvements in pain and disability after 1 and 4 weeks than patients who were positive on the rule and received exercise intervention and patients who were negative on the rule and received manipulation. These results were maintained at the 6-month follow-up. A patient's status on the rule was of little relevance in determining the outcome of patients treated with the exercise intervention. These findings support the hypothesis that

Figure 3. Two-dimensional graphical representation of the 3-way clinical prediction rule × treatment group × time interaction for the Oswestry Disability Questionnaire (ODQ) score ($P < 0.001$).



Lower scores represent less disability.

Table 3. Pairwise Comparisons for the Modified Oswestry Disability Questionnaire Score Change at the 1-Week, 4-Week, and 6-Month Follow-ups*

Comparison	1-Week Modified ODQ Score		4-Week Modified ODQ Score		6-Month Modified ODQ Score	
	Difference (95% CI)	P Value	Difference (95% CI)	P Value	Difference (95% CI)	P Value
Manipulation group vs. exercise group	9.2 (4.4 to 14.1)	<0.001	8.3 (2.4 to 14.2)	0.006	10.1 (4.3 to 15.9)	0.001
Manipulation group (positive on the rule) vs. manipulation group (negative on the rule)	15.0 (8.5 to 21.5)	<0.001	15.2 (7.1 to 23.3)	<0.001	10.3 (2.2 to 18.4)	0.014
Manipulation group (positive on the rule) vs. exercise group (positive on the rule)	20.4 (13.0 to 28.8)	<0.001	14.6 (5.4 to 23.8)	0.003	12.9 (3.5 to 22.3)	0.008
Exercise group (positive on the rule) vs. exercise group (negative on the rule)	-1.9 (4.9 to -8.6)	>0.2	6.5 (-1.8 to 14.8)	0.127	6.8 (-1.5 to 15.2)	0.112

* Higher values represent larger improvements in disability. A similar pattern was observed for pain intensity. ODQ = Oswestry Disability Questionnaire.

the rule is valid for identifying patients with low back pain who are likely to respond rapidly to manipulation.

The purpose of a clinical prediction rule is to improve decision making (12, 42). The rule we studied is designed to enhance decision making about management of patients with low back pain; such guidance has historically been lacking. As seen in the development study (17), in our study a threshold of at least 4 of 5 criteria maximized the positive likelihood ratio for predicting rapid success with manipulation. On the basis of a pretest probability of success of 44% and a positive likelihood ratio of 13.2, a patient positive on the rule and treated with manipulation has a 92% chance of achieving a successful outcome by the end of 1 week (43). Given the low risk related to manipulation of the lumbar spine (44), a shift of this magnitude clearly seems to justify an attempt at spinal manipulation. Even if the lower bound of the 95% CI of 3.4 is presumed to be the point estimate, the post-test probability of success is 73.0%. Conversely, the negative likelihood ratio associated with patients who met fewer than 3 criteria was 0.10. By using the same pretest probability, patients with fewer than 3 criteria would have only a 7% probability of success, indicating the need for alternative treatment. We saw similar accuracy at 4 weeks and 6 months, supporting the prognostic value of the rule at a longer follow-up.

The potential effect of the rule on decision making is also highlighted by the numbers needed to treat. For patients positive on the rule, the number needed to treat for benefit with manipulation was 1.3 at 1 week and 1.9 at 4

weeks. These statistics suggest that only about 2 patients who are positive on the rule need to be treated with manipulation to prevent 1 patient from not achieving a successful outcome after 1 or 4 weeks of treatment. The potential value of such an early, rapid reduction in disability is evident from studies showing that individuals with low back pain who experience persistent disability for as few as 4 weeks are at increased risk for chronic disability and work restrictions (23, 45–47). Because patients with chronic, disabling low back pain account for a disproportionate share of health care expenditures and workers' compensation costs (46), the potential cost savings of an early, effective intervention to prevent even a few individuals from progressing to chronic disability may be considerable. This contention is supported by the 6-month follow-up data. A statistically significantly smaller proportion of patients who were positive on the rule and received manipulation were seeking additional health care for back pain than patients who were negative on the rule but did not receive manipulation. More research on the cost-effectiveness and long-term benefits of the rule is needed to support this hypothesis.

Evidence suggests that patients with early access to physical therapy tend to return to work sooner than when referral is delayed (48). In this study, having symptoms for less than 16 days was the most accurate individual variable in the rule, with a positive likelihood ratio of 4.4 (CI, 2.0 to 9.6). The positive likelihood ratio for this criterion in patients who received the exercise program alone was close to 1.0, suggesting negligible shifts in the probability of a

Table 4. Responses to Questions at 6-Month Follow-up*

Question	Manipulation Group, %			Exercise Group, %		
	Overall (n = 52)	Positive on the Rule (n = 18)	Negative on the Rule (n = 34)	Overall (n = 40)	Positive on the Rule (n = 16)	Negative on the Rule (n = 24)
Have you taken any medications for back pain in the past week?	36.5	27.8	41.2	60.0†	43.8	70.8
Are you presently seeking treatment for back pain?	11.5	11.1	11.8	42.5†	43.8‡	41.7
Have you missed any time at work in the past 6 weeks because of back pain?	9.6	5.6	11.8	25.0†	25.0	25.0

* Numbers are the percentages of patients responding "yes" to the question.

† Significant difference between the overall proportion of patients responding "yes" in the manipulation and exercise groups ($P < 0.05$).

‡ Significant difference between the proportion of patients who were positive on the rule in the manipulation and exercise groups ($P < 0.05$).

successful outcome. However, only 35% of patients (46 of 131) had symptoms for less than 16 days, emphasizing the need for early access to physical therapy intervention in the rehabilitation of patients with low back pain.

Given the 70.2% response rate at the 6-month follow-up, we performed a subsequent “completers only” analysis among patients with all follow-up data available. Similar statistically significant and clinically meaningful findings are observed for the overall multivariate ANOVA and pertinent post hoc comparisons at the 1- and 4-week follow-ups. Compared with the last-observation-carried-forward assumption, just enough statistical power was lost at the 6-month follow-up to result in non-statistically significant differences for the post hoc comparisons. However, the magnitudes of change were maintained, clearly representing improvements that would be clinically meaningful to patients. However, the degree of dropout and the uncertainty in the final outcomes of those who dropped out make this conclusion uncertain.

Our results reinforce the importance of matching individual patients to interventions from which they are likely to benefit. It is highly unreasonable to expect that all patients with low back pain would benefit from any single intervention, including manipulation. This assumption, however, seems to underlie the design of many previous studies that have compared heterogeneous samples, often with equivocal results. Manipulation may benefit a subgroup of patients with low back pain (6), and focusing research on this subgroup could enhance the power of future clinical research. We identified 36% of patients (47 of 131) who were positive on the rule, thus considered likely to benefit from spinal manipulation. The observed effect of manipulation on disability and pain among these patients was statistically significantly greater than the effect of manipulation on the group as a whole.

This rule corresponds to a level II clinical prediction rule (49), and clinicians caring for patients with low back pain can have increased confidence in using the rule to improve decision making. Future research should determine the effect of the rule’s implementation on practice patterns, outcomes of care, and costs (12).

From Wilford Hall Medical Center, Lackland Air Force Base, San Antonio, Texas; University of Pittsburgh, Pittsburgh, Pennsylvania; U.S. Army–Baylor University Graduate Program in Physical Therapy, Fort Sam Houston, Texas; Hill Air Force Base Medical Clinic, Ogden, Utah; and Luke Air Force Base Medical Clinic, Phoenix, Arizona.

Disclaimer: The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the U.S. Air Force or Department of Defense.

Acknowledgments: The authors thank the physical therapy staff at the following sites for their assistance with data collection: Wilford Hall Medical Center, Lackland Air Force Base (AFB); Malcolm Grow Medical Center, Andrews AFB; Wright-Patterson Medical Center, Wright-Patterson AFB; Eglin Hospital, Eglin AFB; Luke Medical Clinic, Luke AFB; Hill Medical Clinic, Hill AFB; F.E. Warren Medical Clinic, F.E.

Warren AFB; and University of Pittsburgh Medical Center Health System’s Centers for Rehab Services.

Grant Support: By the Foundation for Physical Therapy, Inc., and Wilford Hall Medical Center Commander’s Intramural Research Funding Program.

Potential Financial Conflicts of Interest: None disclosed.

Requests for Single Reprints: Maj John D. Childs, PhD, PT, 508 Thurber Drive, Schertz, TX 78154-1146; e-mail, childsjd@sbcglobal.net.

Current author addresses and author contributions are available at www.annals.org.

References

1. Deyo RA, Phillips WR. Low back pain. A primary care challenge. *Spine*. 1996;21:2826-32. [PMID: 9112706]
2. Stewart WF, Ricci JA, Chee E, Morganstein D, Lipton R. Lost productive time and cost due to common pain conditions in the US workforce. *JAMA*. 2003;290:2443-54. [PMID: 14612481]
3. Luo X, Pietrobon R, Sun SX, Liu GG, Hey L. Estimates and patterns of direct health care expenditures among individuals with back pain in the United States. *Spine*. 2004;29:79-86. [PMID: 14699281]
4. Philadelphia Panel evidence-based clinical practice guidelines on selected rehabilitation interventions for low back pain. *Phys Ther*. 2001;81:1641-74. [PMID: 11589642]
5. Shekelle PG, Adams AH, Chassin MR, Hurwitz EL, Brook RH. Spinal manipulation for low-back pain. *Ann Intern Med*. 1992;117:590-8. [PMID: 1388006]
6. Assendelft WJ, Morton SC, Yu EI, Suttrop MJ, Shekelle PG. Spinal manipulative therapy for low back pain. A meta-analysis of effectiveness relative to other therapies. *Ann Intern Med*. 2003;138:871-81. [PMID: 12779297]
7. Cherkin DC, Sherman KJ, Deyo RA, Shekelle PG. A review of the evidence for the effectiveness, safety, and cost of acupuncture, massage therapy, and spinal manipulation for back pain. *Ann Intern Med*. 2003;138:898-906. [PMID: 12779300]
8. Koes BW, van Tulder MW, Ostelo R, Kim Burton A, Waddell G. Clinical guidelines for the management of low back pain in primary care: an international comparison. *Spine*. 2001;26:2504-13; discussion 2513-4. [PMID: 11707719]
9. Borkan JM, Koes B, Reis S, Cherkin DC. A report from the Second International Forum for Primary Care Research on Low Back Pain. Reexamining priorities. *Spine*. 1998;23:1992-6. [PMID: 9779533]
10. Bouter LM, van Tulder MW, Koes BW. Methodologic issues in low back pain research in primary care. *Spine*. 1998;23:2014-20. [PMID: 9779536]
11. van Tulder M, Malmivaara A, Esmail R, Koes B. Exercise therapy for low back pain: a systematic review within the framework of the cochrane collaboration back review group. *Spine*. 2000;25:2784-96. [PMID: 11064524]
12. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users’ guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. *JAMA*. 2000;284:79-84. [PMID: 10872017]
13. Stiell IG, Lesiuk H, Wells GA, McKnight RD, Brison R, Clement C, et al. The Canadian CT Head Rule Study for patients with minor head injury: rationale, objectives, and methodology for phase I (derivation). *Ann Emerg Med*. 2001;38:160-9. [PMID: 11468612]
14. Stiell IG, Wells GA, Vandemheen KL, Clement CM, Lesiuk H, De Maio VJ, et al. The Canadian C-spine rule for radiography in alert and stable trauma patients. *JAMA*. 2001;286:1841-8. [PMID: 11597285]
15. Stiell IG, McKnight RD, Greenberg GH, McDowell I, Nair RC, Wells GA, et al. Implementation of the Ottawa ankle rules. *JAMA*. 1994;271:827-32. [PMID: 8114236]
16. Stiell IG, Wells GA, Hoag RH, Sivilotti ML, Cacciotti TF, Verbeek PR, et al. Implementation of the Ottawa Knee Rule for the use of radiography in acute

- knee injuries. *JAMA*. 1997;278:2075-9. [PMID: 9403421]
17. Flynn T, Fritz J, Whitman J, Wainner R, Magel J, Rendeiro D, et al. A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine*. 2002;27:2835-43. [PMID: 12486357]
 18. Guyatt G, Rennie D, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. Chicago: American Medical Assoc Pr; 2002.
 19. Mann NH 3rd, Brown MD, Hertz DB, Enger I, Tompkins J. Initial-impression diagnosis using low-back pain patient pain drawings. *Spine*. 1993;18:41-53. [PMID: 8434324]
 20. Jensen MP, Turner JA, Romano JM. What is the maximum number of levels needed in pain intensity measurement? *Pain*. 1994;58:387-92. [PMID: 7838588]
 21. Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain*. 1993;52:157-68. [PMID: 8455963]
 22. Jacob T, Baras M, Zeev A, Epstein L. Low back pain: reliability of a set of pain measurement tools. *Arch Phys Med Rehabil*. 2001;82:735-42. [PMID: 11387576]
 23. Fritz JM, George SZ. Identifying psychosocial variables in patients with acute work-related low back pain: the importance of fear-avoidance beliefs. *Phys Ther*. 2002;82:973-83. [PMID: 12350212]
 24. Crombez G, Vlaeyen JW, Heuts PH, Lysens R. Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. *Pain*. 1999;80:329-39. [PMID: 10204746]
 25. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine*. 2000;25:2940-52; discussion 2952. [PMID: 11074683]
 26. Fritz JM, Irrgang JJ. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Phys Ther*. 2001;81:776-88. [PMID: 11175676]
 27. Waddell G, Somerville D, Henderson I, Newton M. Objective clinical evaluation of physical impairment in chronic low back pain. *Spine*. 1992;17:617-28. [PMID: 1308095]
 28. Freburger JK, Riddle DL. Using published evidence to guide the examination of the sacroiliac joint region. *Phys Ther*. 2001;81:1135-43. [PMID: 11319939]
 29. Malmivaara A, Häkkinen U, Aro T, Heinrichs ML, Koskenniemi L, Kusosma E, et al. The treatment of acute low back pain—bed rest, exercises, or ordinary activity? *N Engl J Med*. 1995;332:351-5. [PMID: 7823996]
 30. Childs JD, Fritz JM, Piva SR, Erhard RE. Clinical decision making in the identification of patients likely to benefit from spinal manipulation: a traditional versus an evidence-based approach. *J Orthop Sports Phys Ther*. 2003;33:259-72. [PMID: 12775000]
 31. Bigos S, Bowyer O, Braen G, Brown K, Deyo R, Haldeman S, et al. *Acute Lower Back Problems in Adults*. Rockville, MD: Agency for Health Care Policy and Research; 1994. AHCPR publication no. 95-0642.
 32. van Tulder MW, Koes BW, Bouter LM. Conservative treatment of acute and chronic nonspecific low back pain. A systematic review of randomized controlled trials of the most common interventions. *Spine*. 1997;22:2128-56. [PMID: 9322325]
 33. McGill SM. Low back exercises: evidence for improving exercise regimens. *Phys Ther*. 1998;78:754-65. [PMID: 9672547]
 34. Fritz JM, Hicks GE. Exercise protocols for low back pain. In: Liemohn WP, ed. *Exercise Prescription and the Back*. New York: McGraw-Hill; 2001:167-82.
 35. Erhard RE, Delitto A, Cibulka MT. Relative effectiveness of an extension program and a combined program of manipulation and flexion and extension exercises in patients with acute low back syndrome. *Phys Ther*. 1994;74:1093-100. [PMID: 7991650]
 36. Fritz JM, George S. The use of a classification approach to identify subgroups of patients with acute low back pain. Interrater reliability and short-term treatment outcomes. *Spine*. 2000;25:106-14. [PMID: 10647168]
 37. Glass GV, Hopkins KD. *Statistical Methods in Education and Psychology*. 3rd ed. Boston: Allyn and Bacon; 1996.
 38. Fritz JM, Wainner RS. Examining diagnostic tests: an evidence-based perspective. *Phys Ther*. 2001;81:1546-64. [PMID: 11688591]
 39. van Tulder MW, Scholten RJ, Koes BW, Deyo RA. Nonsteroidal anti-inflammatory drugs for low back pain: a systematic review within the framework of the Cochrane Collaboration Back Review Group. *Spine*. 2000;25:2501-13. [PMID: 11013503]
 40. Hagen KB, Hilde G, Jamtvedt G, Winnem MF. The cochrane review of advice to stay active as a single treatment for low back pain and sciatica. *Spine*. 2002;27:1736-41. [PMID: 12195064]
 41. van Tulder MW, Malmivaara A, Esmail R, Koes BW. Exercise therapy for low back pain. *Cochrane Database Syst Rev*. 2000:CD000335. [PMID: 10796344]
 42. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. *JAMA*. 1997;277:488-94. [PMID: 9020274]
 43. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence-Based Medicine: How to Practice and Teach EBM*. New York: Churchill Livingstone; 2000.
 44. Jaeschke R, Guyatt GH, Sackett DL. *Users' guides to the medical literature*. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA*. 1994;271:703-7. [PMID: 8309035]
 45. Frank JW, Brooker AS, DeMaio SE, Kerr MS, Maetzel A, Shannon HS, et al. Disability resulting from occupational low back pain. Part II: What do we know about secondary prevention? A review of the scientific evidence on prevention after disability begins. *Spine*. 1996;21:2918-29. [PMID: 9112717]
 46. Hashemi L, Webster BS, Clancy EA. Trends in disability duration and cost of workers' compensation low back pain claims (1988-1996). *J Occup Environ Med*. 1998;40:1110-9. [PMID: 9871888]
 47. Hiebert R, Skovron ML, Nordin M, Crane M. Work restrictions and outcome of nonspecific low back pain. *Spine*. 2003;28:722-8. [PMID: 12671363]
 48. Ehrmann-Feldman D, Rossignol M, Abenheim L, Gobeille D. Physician referral to physical therapy in a cohort of workers compensated for low back pain. *Phys Ther*. 1996;76:150-6; discussion 156-7. [PMID: 8592718]
 49. McGinn T, Guyatt G, Wyer P, Naylor CD, Stiell I. Diagnosis: Clinical Prediction Rule. In: Guyatt G, Rennie D, eds. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*. Chicago: American Medical Assoc Pr; 2002:471-83.
 50. Chan CW, Goldman S, Ilstrup DM, Kunselman AR, O'Neill PI. The pain drawing and Waddell's nonorganic physical signs in chronic low-back pain. *Spine*. 1993;18:1717-22. [PMID: 8235853]
 51. Udén A, Åström M, Bergenudd H. Pain drawings in chronic back pain. *Spine*. 1988;13:389-92. [PMID: 2970123]
 52. Werneke MW, Harris DE, Lichter RL. Clinical effectiveness of behavioral signs for screening chronic low-back pain patients in a work-oriented physical rehabilitation program. *Spine*. 1993;18:2412-8. [PMID: 8303442]
 53. Maitland GD. *Vertebral Manipulation*. 5th ed. London: Butterworth-Heinemann; 1986.
 54. Maher CG, Latimer J, Adams R. An investigation of the reliability and validity of posteroanterior spinal stiffness judgments made using a reference-based protocol. *Phys Ther*. 1998;78:829-37. [PMID: 9711208]
 55. Binkley J, Stratford PW, Gill C. Interrater reliability of lumbar accessory motion mobility testing. *Phys Ther*. 1995;75:786-92; discussion 793-5. [PMID: 7659738]
 56. Gonnella C, Paris SV, Kutner M. Reliability in evaluating passive intervertebral motion. *Phys Ther*. 1982;62:436-44. [PMID: 7063535]
 57. Maher C, Adams R. Reliability of pain and stiffness assessments in clinical manual lumbar spine examination. *Phys Ther*. 1994;74:801-9; discussion 809-11. [PMID: 8066107]
 58. Lundberg G, Gerdle B. The relationships between spinal sagittal configuration, joint mobility, general low back mobility and segmental mobility in female homecare personnel. *Scand J Rehabil Med*. 1999;31:197-206. [PMID: 10599896]
 59. Strender LE, Sjöblom A, Sundell K, Ludwig R, Taube A. Interexaminer reliability in physical examination of patients with low back pain. *Spine*. 1997;22:814-20. [PMID: 9106324]
 60. Ellison JB, Rose SJ, Sahrman SA. Patterns of hip rotation range of motion: a comparison between healthy subjects and patients with low back pain. *Phys Ther*. 1990;70:537-41. [PMID: 2144050]

APPENDIX 1: METHODS FOR ASSESSING A PATIENT'S STATUS ON EACH CRITERION IN THE SPINAL MANIPULATION CLINICAL PREDICTION RULE

1. Duration of Current Episode of Symptoms Less than 16 Days

Patients are asked to report the number of days since the onset of their current episode of low back pain.

2. Location of Symptoms Not Extending Distal to the Knee

A body diagram is used to assess the distribution of symptoms (19, 50, 51). We categorize the location of symptoms as being in the back, buttock, thigh, or leg (distal to knee) by using the method described by Werneke and colleagues (52), who found high inter-rater reliability ($\kappa = 0.96$).

3. Score on the FABQ Work Subscale Less than 19 Points

The FABQ (21) is subdivided into 2 subscales, a 5-item physical activity subscale (questions 1 to 5) and a 16-item work subscale (questions 6 to 16). Decision making using the rule requires only the FABQ work subscale score. However, all items on the questionnaire should be completed since they were included when the psychometric properties of the instrument were established. Each item is scored from 0 to 6; however, not all items within each subscale contribute to the score. Four items (items 2, 3, 4, and 5) are scored for the FABQ physical activity subscale, and 7 items (items 6, 7, 9, 10, 11, 12, and 15) are scored for the FABQ work subscale. Each scored item within a particular subscale is summed; thus, possible scores range from 0 to 42 and 0 to 28 for the FABQ work and FABQ physical activity subscales, respectively. Higher scores represent increased fear-avoidance beliefs.

4. At Least 1 Lumbar Spine Segment Judged To Be Hypomobile

Segmental mobility of the lumbar spine is tested with the patient prone and the neck in neutral rotation. Testing is performed over the spinous processes of the vertebrae (53, 54). The examiner stands at the head or side of the table and places the hypothenar eminence of the hand (that is, the pisiform bone) over the spinous process of the segment to be tested. With the elbow and wrist extended, the examiner applies a gentle but firm, anteriorly directed pressure on the spinous process. The stiffness at each segment is judged as normal, hypomobile, or hypermobile. The examiner interpreted whether a segment is hypomobile on the basis of the examiner's anticipation of what normal mobility would feel like at that level and compared with the mobility detected in the segment above and below. Some authors have reported poor inter-rater reliability for judgments of spinal segmental mobility on scales with 7 to 11 levels of judgments (55–57). Studies using mobility judgments similar to those in our study have reported adequate inter-rater reliability ($\kappa = 0.40$ to 0.68) (58, 59).

5. At Least 1 Hip with More than 35 Degrees of Internal Rotation Range of Motion

Hip range of motion is tested bilaterally with the patient lying prone and with the cervical spine at the midline. The examiner places the leg opposite that to be measured in approximately 30 degrees of hip abduction to enable the tested hip to be freely moved. The lower extremity of the side to be tested is kept in line with the body, and the knee on that side is flexed to 90 degrees. A gravity inclinometer is placed on the distal aspect of the fibula in line with the bone. Internal rotation is measured at the point in which the pelvis first begins to move. Ellison and colleagues (60) reported excellent inter-rater reliability with these procedures (intraclass correlation coefficients, 0.95 to 0.97).

APPENDIX 2: PROCEDURES USED TO PERFORM THE SPINAL MANIPULATION INTERVENTION

All patients received the same technique. The patient was supine. The physical therapist stood opposite the side to be manipulated and moved the patient into side-bending toward the side to be manipulated. The patient was asked to interlock the fingers behind the head. The physical therapist then rotated the patient and delivered a quick thrust to the pelvis in a posterior and inferior direction (Figure 1). The side to be manipulated was the more symptomatic side on the basis of the patient's report. If the patient could not specify a side, the physical therapist selected a side to be manipulated. If a cavitation (that is, a "pop") occurred, the physical therapist instructed the patient in the range-of-motion exercise. If no cavitation was produced, the patient was repositioned and the manipulation was attempted again. A maximum of 2 attempts per side was permitted. If no cavitation was produced after the fourth attempt, the physical therapist proceeded to instruct the patient in the range-of-motion exercise. Patients were instructed to perform 10 repetitions of the range-of-motion exercise in the clinic and 10 repetitions 3 to 4 times daily on the days that they did not attend physical therapy. Beginning with the third session, patients in the manipulation group completed the same exercise program as patients in the exercise group.

Current Author Addresses: Dr. Childs: 508 Thurber Drive, Schertz, TX 78154.

Dr. Fritz: Department of Physical Therapy, University of Utah, 520 Wakara Way, Salt Lake City, UT 84108.

Dr. Flynn: Department of Physical Therapy, Regis University, 3333 Regis Boulevard, G-4, Denver, CO 80221-1099.

Drs. Irrgang and Delitto: Department of Physical Therapy, University of Pittsburgh, 6035 Forbes Tower, Pittsburgh, PA 15260.

Mr. Majkowski: Physical Therapy Service, 3851 Roger Brooke Drive, Fort Sam Houston, TX 78234.

Mr. Johnson: 2602 Blue Rock Drive, Beavercreek, OH 45434.

Author Contributions: Conception and design: J.D. Childs, J.M. Fritz, T.W. Flynn, J.J. Irrgang, A. Delitto. Administrative, technical, or logistic support: K.K. Johnson, G.R. Majkowski. Collection and assembly of data: J.D. Childs, K.K. Johnson, G.R. Majkowski.

Analysis and interpretation of the data: J.D. Childs, J.M. Fritz, T.W. Flynn, J.J. Irrgang.
Drafting of the article: J.D. Childs, J.M. Fritz, A. Delitto.
Critical revision of the article for important intellectual content: J.D. Childs, J.M. Fritz, T.W. Flynn, K.K. Johnson, A. Delitto.

Final approval of the article: J.D. Childs, J.M. Fritz, T.W. Flynn, J.J. Irrgang, K.K. Johnson, G.R. Majkowski, A. Delitto.
Provision of study materials or patients: K.K. Johnson, G.R. Majkowski.
Statistical expertise: J.J. Irrgang.
Obtaining of funding: J.D. Childs, J.M. Fritz.

Treatments for Back Pain: Can We Get Past Trivial Effects?

The history of treatments for back and neck pain is generally one of small increments in benefit. For patients with acute pain, the natural history is favorable, and it is hard to prove that most treatments result in greater improvement than the nonspecific effects of natural history, placebo, and regression to the mean. The treatments that appear better than placebo, such as nonsteroidal anti-inflammatory drugs, muscle relaxants, and spinal manipulation, generally have modest effects (1–3). For chronic pain, the natural history is less favorable, but it is similarly difficult to prove that most treatments are better than nonspecific effects alone. The literature is replete with conflicting results, modest effects, and weak studies. The treatments that may be better than placebo, such as exercise programs, tricyclic antidepressants, and cognitive behavioral therapy, often require substantial patient commitment or lifestyle changes. In both clinical practice and clinical trials, many patients decline these treatments or find adherence difficult.

Furthermore, drug side effects and surgical risks attend many conventional treatments. Patients say that doctors spend too little time, don't understand their problems, or are dismissive. The combination of desperate patients, clinicians eager to offer relief, advocacy for favored approaches, lack of definitive end points (death or cure), and misleading nonspecific effects has promoted wide adoption of treatments that have ultimately been discredited. These include bed rest, lumbar traction, sacroiliac fusion, and others, all of dubious efficacy.

It should be no surprise, then, that patients turn increasingly to complementary and alternative treatments for symptoms like back and neck pain. Spinal manipulation, acupuncture, and massage therapy are the most commonly used alternative treatments for these conditions. Such treatments offer the allure of glowing patient testimonials, enthusiastic practitioners, and reasonable safety. But if public and private insurance plans are to cover these treatments, it is only fair to hold the treatments to the same scientific standards we expect of conventional care.

Two articles in this issue address "alternative" treatments for spinal disorders: manipulation for acute low back pain and acupuncture for chronic neck pain. The study by Childs and colleagues (4) tests the validity of a prediction rule for selecting patients likely to benefit from spinal manipulation, and the study by White and associates (5) tests the efficacy of "western" acupuncture. In brief, spinal manipulation for acute low back pain, and the prediction rule, looked good. Acupuncture for chronic neck pain was disappointing; the average effect on pain was statistically significant but failed to achieve a prespecified threshold of clinical importance. Functional outcomes were not improved over placebo.

In the study by Childs and colleagues, spinal manipu-

lation seemed to offer a "slam-dunk" effect if patients met certain criteria. The study was a well-designed randomized trial and indicated that patients who met the criteria had a much greater effect from the intervention than those who did not. Furthermore, those who met the criteria benefited more from manipulation than from an exercise intervention. Some caveats are in order, however.

First, Childs and colleagues may have compared manipulation with an ineffective therapy, even though a decade-old guideline implied otherwise. As the authors point out, the guideline for acute back pain developed by the Agency for Health Care Policy and Research made a weak recommendation for exercise (6). However, several randomized trials in the past decade suggest that exercise is an ineffective treatment for *acute* back pain, as opposed to chronic back pain, for which it is effective (7). A meta-analysis suggested that manipulation is consistently more effective than discredited treatments, but no more effective than conventional treatments of demonstrated efficacy (3). The study by Childs and colleagues may in effect be another positive comparison of manipulation with placebo therapy, rather than with an effective alternative.

Generalizability is also problematic. Childs and colleagues' study was performed mainly in military clinics, where the clientele may differ from most primary care settings. The manipulation was performed by "conventional" physical therapists. Because manipulative techniques can vary widely, it remains unclear whether the technique or its demonstrated efficacy is better than, worse than, or equivalent to that of chiropractors or osteopathic physicians. In my state (Washington) and some others, physical therapists are legally proscribed from performing manipulation.

In White and associates' study of acupuncture, a single therapist provided all treatment, using "western" acupuncture techniques. Patients with litigation or disability claims were excluded. Whether these results would occur with other practitioners, "traditional Chinese" acupuncture, or patients involved in disability claims is not known.

This study is one of several that has found acupuncture to be ineffective for spinal disorders (8, 9). One possible reason is that patients may not have expected much benefit from acupuncture. In a randomized trial comparing massage with acupuncture, patients who received massage did better if they expected it to be effective than if they expected acupuncture would be more effective. Among the patients randomly assigned to acupuncture, those who expected acupuncture to be most helpful had a much greater improvement than those who expected massage to be more effective (10). The magnitude of the differences based on patient expectations was easily as large as the treatment effect described in many "positive" trials. Expectations may have more influence in trials of physical treatments, where blinding is often impossible, than in successfully blinded

drug trials. Previous expectations could have blunted the apparent benefit of acupuncture in the trial by White and associates if most patients in both groups thought acupuncture would be ineffective. Similarly, expectations might have amplified the apparent benefit of manipulation in Childs and colleagues' study if most patients in both groups expected manipulation to be helpful.

What should we conclude from these 2 trials? At the least, Childs and colleagues' study of manipulation suggests a useful predictive model that may improve the design of future trials of spinal manipulation. The study also is an excellent model for studying similar classification and prediction rules. It is gratifying to see a patient "classification" scheme validated so convincingly. This approach may help us to improve on the often trivial benefits of many back pain treatments by matching patients with the treatments they are most likely to respond to.

White and associates' acupuncture study adds to a growing literature suggesting that acupuncture has little effect in patients with spinal disorders, as 2 recent meta-analyses have shown (8, 9). Perhaps acupuncture will prove less effective in western societies than in Asian cultures in part because western patients have lower expectations of benefit than Asian patients have.

On the basis of current knowledge, where do manipulation and acupuncture fit in the management of spinal disorders? Manipulation seems to be a reasonable nonpharmacologic option for treating low back pain because it is probably equivalent in efficacy to most conventional treatments. The decision rule tested by Childs and colleagues may help to identify the patients most likely to benefit. The efficacy and safety of manipulation for neck pain are less certain. The efficacy of acupuncture for either low back or neck pain has been disappointing in randomized trials so far. It may be best reserved for patients who respond poorly to other treatments and who have a strong conviction that it will help.

Both trials in this issue suggest a need for better understanding of specific patient expectations, their influences on treatment outcomes, and their implications for recommending specific treatments. Both studies also point to the need for expanded research on common disabling musculoskeletal conditions. Where are the large trials involving dozens of practitioners and thousands of patients that are commonplace in cardiovascular disease and oncology and so scarce in musculoskeletal diseases? The nearly ubiquitous musculoskeletal conditions that "merely" disable millions deserve equal attention and scientifically rigorous study.

Richard A. Deyo, MD, MPH
University of Washington
Seattle, WA 98195

Disclaimer: The opinions and conclusions herein are those of the author and not necessarily those of the National Institute for Arthritis, Musculoskeletal, and Skin Diseases.

Grant Support: By the National Institute for Arthritis, Musculoskeletal, and Skin Diseases (No. P60 AR48093).

Potential Financial Conflicts of Interest: None disclosed.

Requests for Single Reprints: Richard A. Deyo, MD, MPH, Box 357183, University of Washington, Seattle, WA 98195; e-mail, deyo@u.washington.edu.

Ann Intern Med. 2004;141:957-958.

References

1. van Tulder MW, Scholten RJ, Koes BW, Deyo RA. Non-steroidal anti-inflammatory drugs for low back pain. *Cochrane Database Syst Rev.* 2000; CD000396. [PMID: 10796356]
2. van Tulder MW, Touray T, Furlan AD, Solway S, Bouter LM. Muscle relaxants for non-specific low back pain. *Cochrane Database Syst Rev.* 2003; CD004252. [PMID: 12804507]
3. Assendelft WJ, Morton SC, Yu EI, Suttrop MJ, Shekelle PG. Spinal manipulative therapy for low back pain. A meta-analysis of effectiveness relative to other therapies. *Ann Intern Med.* 2003;138:871-81. [PMID: 12779297]
4. Childs JD, Fritz JM, Flynn TW, Irrgang JJ, Johnson KK, Majkowski GR, et al. A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Ann Intern Med.* 2004;141: 920-8.
5. White P, Lewith G, Prescott P, Conway J. Acupuncture versus placebo for the treatment of chronic mechanical neck pain. A randomized, controlled trial. *Ann Intern Med.* 2004;141:911-9.
6. Bigos S, Bowyer O, Braen G, et al. Acute Low Back Problems in Adults. Clinical Practice Guideline No. 14. Rockville, MD: Agency for Health Care Policy and Research, Public Health Service, U.S. Department of Health and Human Services; 1994. AHCPR publication no. 95-0642.
7. van Tulder MW, Malmivaara A, Esmail R, Koes BW. Exercise therapy for low back pain. *Cochrane Database Syst Rev.* 2000;CD000335. [PMID: 10796344]
8. van Tulder MW, Cherkin DC, Berman B, Lao L, Koes BW. Acupuncture for low-back pain. *Cochrane Database of Systematic Reviews.* 1999, Issue 2. CD001351. DOI: 10.1002/14651858.CD001351.
9. Cherkin DC, Sherman KJ, Deyo RA, Shekelle PG. A review of the evidence for the effectiveness, safety, and cost of acupuncture, massage therapy, and spinal manipulation for back pain. *Ann Intern Med.* 2003;138:898-906. [PMID: 12779300]
10. Kalaoukalani D, Cherkin DC, Sherman KJ, Koepsell TD, Deyo RA. Lessons from a trial of acupuncture and massage for low back pain: patient expectations and treatment effects. *Spine.* 2001;26:1418-24. [PMID: 11458142]

© 2004 American College of Physicians