

Maruti Ram Gudavalli
Jerrilyn A. Cambron
Marion McGregor
James Jedlicka
Michael Keenum
Alexander J. Ghanayem
Avinash G. Patwardhan

A randomized clinical trial and subgroup analysis to compare flexion–distraction with active exercise for chronic low back pain

Received: 8 December 2004
Revised: 18 August 2005
Accepted: 1 November 2005
Published online: 8 December 2005
© Springer-Verlag 2005

M. R. Gudavalli (✉)
Palmer College of Chiropractic, Research,
Davenport IA, USA
E-mail: gudavalli_r@palmer.edu
Tel.: +1-563-8845260
Fax: +1-563-8845227

M. R. Gudavalli · J. A. Cambron
J. Jedlicka
National University of Health Sciences,
Research, Lombard IL, USA

M. Keenum
Orthosport Physical Therapy, Research,
Forest Park IL, USA

A. J. Ghanayem · A. G. Patwardhan
Orthopedic Surgery, Loyola University
Stritch School of Medicine,
Maywood IL, USA

A. J. Ghanayem · A. G. Patwardhan
Rehabilitation Research, Edward Hines,
Jr. VA Hospital, Hines IL, USA

M. McGregor
Research Consultant,
Richardson TX, USA

Abstract Many clinical trials on chiropractic management of low back pain have neglected to include specific forms of care. This study compared two well-defined treatment protocols. The objective was to compare the outcome of flexion–distraction (FD) procedures performed by chiropractors with an active trunk exercise protocol (ATEP) performed by physical therapists. A randomized clinical trial study design was used. Subjects, 18 years of age and older, with a primary complaint of low back pain (> 3 months) were recruited. A 100 mm visual analogue scale (VAS) for perceived pain, the Roland Morris (RM) Questionnaire for low back function, and the SF-36 for overall health status served as primary outcome measures. Subjects were randomly allocated to receive either FD or ATEP. The FD intervention consisted of the application of flexion and traction applied to specific regions in the low back, with the aid of a specially designed manipulation table. The ATEP intervention included stabilizing and flexibility exercises, the use of modalities, and cardiovascular training. A total of 235 subjects met the inclusion/exclusion criteria and signed the informed consent. Of these, 123 were randomly allocated to FD and 112 to the ATEP. Study patients perceived significantly less pain and better function after intervention, regard-

less of which group they were allocated to ($P < 0.01$). Subjects randomly allocated to the flexion–distraction group had significantly greater relief from pain than those allocated to the exercise program ($P = 0.01$). Subgroup analysis indicated that subjects categorized as chronic, with moderate to severe symptoms, improved most with the flexion–distraction protocol. Subjects categorized with recurrent pain and moderate to severe symptoms improved most with the exercise program. Patients with radiculopathy did significantly better with FD. There were no significant differences between groups on the Roland Morris and SF-36 outcome measures. Overall, flexion–distraction provided more pain relief than active exercise; however, these results varied based on stratification of patients with and without radiculopathy and with and without recurrent symptoms. The subgroup analysis provides a possible explanation for contrasting results among randomized clinical trials of chronic low back pain treatments and these results also provide guidance for future work in the treatment of chronic low back pain.

Keywords Low back pain · Chiropractic · Physical therapy · Chronic pain · Randomized clinical trial

Introduction

Despite years of work, low back pain remains a substantial healthcare burden. It is the second most common symptom leading patients to seek medical care [8], and claims for this disorder account for the highest proportion of costs of ongoing welfare benefits under the Social Security Disability Income program for workers less than 45 years of age [20].

Although many clinical research studies have been conducted, including investigations comparing chiropractic treatment and exercise to each other and to other modes of care, results have been disappointing and contradictory [2, 18, 22, 23, 32]. Possible reasons for conflicting evidence among trials surveyed in meta-analyses and reviews include:

1. Diversity of definitions for “chronic” pain used [1]
2. Lack of understanding of potential subgroups in the clinical population [4, 9, 18, 33, 34]
3. Poor specification of treatments [22, 23], resulting in different interventions being compared as if they were the same.

In a majority of chiropractic intervention-based studies, manipulative interventions were poorly defined [23]; however, most appear to have been high-velocity, low-amplitude (HVLA) thrusts.

Flexion–distraction (developed by James Cox, DC, DACBR) [7] is one chiropractic procedure that differs from HVLA in that it is a slow manual traction and mobilization rather than a higher load and speed technique (Fig. 1). According to a national survey, 53–58% of chiropractors utilize flexion–distraction in the management of low back pain [5, 6].

The purpose of this study was to compare the effectiveness of a flexion–distraction (FD) technique admin-

istered by chiropractors and an active trunk exercise program (ATEP) administered by physical therapists for chronic low back pain. Chronic low back pain, for this trial, was defined as back pain greater than 3 months in duration and is consistent with authors such as Helmhout et al. [15], Vollenbroek-Hutton et al. [34] and Katz et al. [17]. The intent was to use a randomized controlled trial to evaluate differences between groups and to include a subgroup analysis to address conflicting evidence in previous studies and guide future research. Turk [33] suggested that patients with a variety of chronic pain syndromes may be clinically hampered by what he referred to as “the patient homogeneity myth”. In other words, the belief that patients with the same diagnosis are sufficiently similar in important variables related to treatment. While Turk emphasized the importance of matching patient subgroups to treatment based on behavioral characteristics, one aim of this study included an evaluation of subgroups based on physical characteristics (radiculopathy vs. no radiculopathy) and history (recurrent pain vs. continuous chronic pain and pain severity). The study and methods were approved by the National University of Health Sciences institutional review board (IRB), approval number H-9701.

Methods

Participants

Table 1 provides the inclusion/exclusion criteria. Subjects were at least 18 years old, with a primary complaint of low back pain for more than 3 months, and had no contraindications to manual therapy (e.g., fracture, dislocation, localized acute infection). Subjects meeting

Fig. 1 Photograph showing a flexion–distraction treatment



Table 1 Inclusion and exclusion criteria

Patients were included in the study if they met the following criteria
1. Primary complaint of low back pain from L1 to SI joint inclusive
2. Duration of low back pain for more than 3 months
3. Palpatory tenderness over one or more lumbar Zygapophyseal joints
4. Willing to forego narcotic use during treatment phase of study
5. Willing to forego NSAID use and/or muscle relaxant use for 24 h prior to baseline or outcome measure assessment
Patients were excluded if they met the following criteria
1. Age less than 18 years
2. Evidence of central nervous system disease
3. Contraindication to manual therapy
(a) Severe osteoporosis
(b) Lumbar fracture
(c) Systemic disease potentially affecting the musculoskeletal system
(d) Failed fusion surgery with unstable components
(e) Inability to undergo physical therapy or flexion–distraction therapy for any other reason
4. Psychiatric illnesses or lack of cognitive abilities that would potentially modify true responses on primary outcomes
5. Current and known substance abuse
6. Not fluent and/or illiterate in the English language
7. Morbidly obese (40% over ideal body weight)
8. Pregnant
9. Currently receiving care for low back pain from any other provider
10. Treated by chiropractor or physical therapist in the past 6 months
11. Not willing to forego care for low back pain at any clinic other than those involved in this study during the 4-week treatment phase
12. New York Heart Association Classification of grade III or IV (marked limitation or inability to carry on any physical activity without discomfort)

the criteria viewed a 3 min video demonstrating treatments and assessments and were presented with an IRB-approved informed consent.

Consecutive new patients with chronic low back pain were initially recruited from two chiropractic clinics and two allopathic (orthopedic) clinics in a Chicago suburb. The physical therapy center was not a primary contact facility and was therefore not included in the recruitment process. Thus both chiropractic and medical physicians were involved in the study, although the study was unsuccessful in recruiting eligible subjects from the medical clinics. Reasons for ineligibility are described in Table 2. Additional recruitment efforts included media advertising such as radio commercials, newspaper advertisements, press releases, cable television advertisements, local posters, and a local electronic sign advertisement. The percentages of patients recruited using each of the methods are: radio advertisements (57%), newspaper advertisements (8%), posters (7%), chiropractic clinic patients (20%), others (e.g., word of mouth) (8%), and orthopedic surgery clinics (0%). Patients received \$75.00 compensation at the completion

Table 2 Reasons for ineligibility

Reason for ineligibility ^a	Number of subjects (N = 1,941)
No current low back pain	1,146
Duration of low back pain < 3 months	324
Current or past medical history of severe systemic disease	127
No palpatory tenderness over lumbar facet joints	101
Less than 18 years of age	85
Current lumbar fracture or instability of the spine	45
Marked limitation or inability to carry on physical activity without cardiac discomfort	42
Morbidly obese	33
Psychiatric illness or lack of cognitive ability	32
Not interested in study participation	28
Current history of severe osteoporosis	27
Evidence of central nervous system disease	23
Treated by a chiropractor in the past 6 months	22
Current and known substance abuse	19
Not fluent or literate in English language	17
Currently receiving care for LBP from other provider, therapist, or physician	15
Treated by a physical therapist in the past 6 months	15
Currently pregnant	1
Not willing to sign the informed consent	77

^aSubjects may be excluded for more than one reason

of active care and an additional \$75.00 upon completion of the 1-year follow-up.

For administrative ease, all data were collected at the chiropractic clinics only. Research personnel were trained in phone screening, administering informed consent, scheduling procedures, and collecting outcome data. Treatment was administered by providers at the chiropractic clinics and at local physical therapy centers.

Interventions

Two treatment arms provided the basis of data collection: a series of flexion–distraction procedures administered by licensed chiropractors who had been certified in this technique after 24 h of training and the completion of two examinations [6, 7] and an ATEP administered by licensed physical therapists.

The FD technique was performed on a specially constructed table with a moveable headpiece, a stationary thoraco-lumbar piece, and a moveable lower extremity piece (Fig. 1). With the subject lying prone, the clinician placed one hand over the lumbar region at the level of interest and used the other hand to flex, laterally flex, and/or rotate the lower extremity section of the table. FD consisted of two biomechanical components. The first component was a series of

traction procedures using the flexion range of motion directed at a specified joint level. The motion from the traction procedure resulted in opening of the posterior joint space and a consequent reduction in intradiscal pressure [7, 12–14]. The second component was a series of mobilization procedures using a possible combination of ranges of motion targeted again at a specific joint level. Most patients moved from the traction component to the mobilization component within 4 weeks of care.

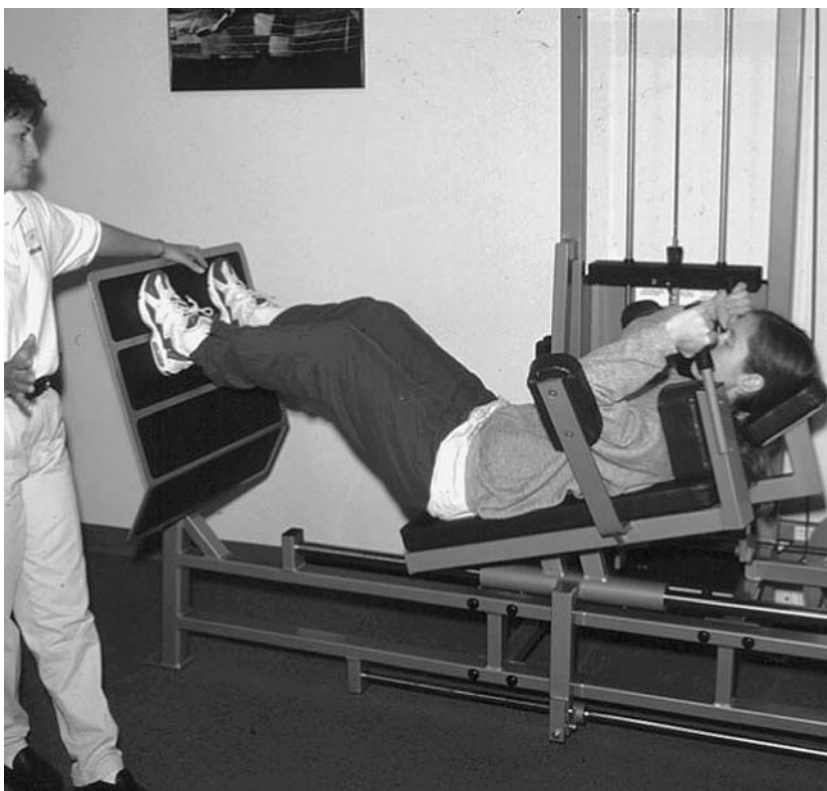
Tolerance testing preceded each FD application. A single flexion–distraction treatment lasted between 3 and 6 min. For patients with radiculopathy, the traction procedure in flexion was used. The number of repetitions was determined by symptom severity. Three sets of repetitions were given at each visit. Each repetition was held for 4 s and a maximum of five repetitions were allowed per set. For patients without radicular symptoms, the mobilization procedures were used. Along with flexion, all physiologic ranges of motion were allowed. Repetitions in all motions except circumduction were held for 2 s each. For circumduction, a single repetition lasted 4 s. Again, the number of repetitions was determined by the severity of symptoms with a maximum of 15 for each set. Meetings were held to insure consistency of technique application.

For each study patient at each visit, all clinically relevant vertebral levels from the lower thoracic spine through the L5/S1 level were treated. These were

determined through palpation of the lower thoracic and lumbar spine and through other signs and symptoms associated with each case. Both motion palpation of the facet joints and static palpation (application of pressure to tolerance directly over the facets [24]) were used to confirm facet joint tenderness. The FD group also received modalities such as ultrasound and cryotherapy. The FD intervention was administered by chiropractors with postgraduate certification in this technique. Application of treatment protocols was assessed and consistency between providers was confirmed by file review and meetings for each profession.

ATEP was administered by licensed physical therapists and consisted of flexion or extension exercises, weight training, flexibility exercises, and cardiovascular exercises dependent on patient symptoms. The aim of the program was to strengthen the muscles surrounding the spine and increase flexibility (Fig. 2). Methods used to develop stabilizing exercises were consistent with those of O’Sullivan and colleagues [28, 29]. The therapists in the study met as a group to choose the specific exercise regime for study purposes and met monthly to reinforce treatment consistency. Biomechanically, the ATEP did not concentrate on a specific joint level but sought to impact the lumbar spine as a whole. It consisted of four phases. During phase one, dependent on clinically relevant symptoms, study subjects received flexion or extension exercises, flexibility exercises,

Fig. 2 Photograph showing an active exercise treatment



modalities such as ultrasound and cryotherapy, and an individualized cardiovascular exercise program. Study subjects with symptoms below the knee received McKenzie extension exercises during phase one. For phase two, upper and lower extremity weight training was added. At phase three, lumbar extension training was included and by phase four, a second cardiovascular exercise and an increase in weight training (to tolerance) were added. Treatment times for the physical therapy group lasted from 30 to 45 min.

Study participants in both treatment groups were seen 2–4 times per week, at the discretion of the treatment provider, for 4 weeks.

Objectives

The two objectives of this investigation were to compare the effectiveness of the FD intervention with the ATEP intervention and to use the study data to consider feasibility issues that may have impacted previous investigations. Specifically, the primary null hypotheses were:

1. Perception of pain as measured by a 100 mm visual analogue scale (VAS) will not differ between groups after 4 weeks of either FD or ATEP.
2. Function as measured by the Roland Morris (RM) Questionnaire will not differ between groups under these study conditions.
3. Overall health status as measured by the SF-36 will not differ between groups.

Feasibility issues involved the logistics of multidisciplinary research conducted at multiple sites, guidance for future trials with respect to sample size estimates, and subgroup analysis for better quality inclusion/exclusion criteria to match subjects with contrasting interventions. As stabilization exercises were included in ATEP, one subgroup division was based on recurrent versus chronic pain. Pain severity at study onset was also used to create divisions since treatment response may differ as a result of initial conditions such as this. Finally, subgroup division according to evidence of radiculopathy (yes/no) was included, since the prognosis for patients with radiculopathy may be different from those without.

Outcomes

Three primary outcome measures were used in this investigation, a 100 mm VAS for perceived pain [16, 21], the RM to measure function [11, 30], and the SF-36 as a representation of health status [3, 11, 21]. Primary outcome measures were taken at baseline and again upon completion of the 4-week intervention period. The results of this initial 4-week time frame are presented here.

Outcome assessments occurred on the day of randomization and at the debriefing visit, which occurred within 48 h of the last treatment. At the end of the investigation, subjects were surveyed about their level of satisfaction for the overall study period. Standard questions regarding overall satisfaction, future use, and recommendation for others were created and are consistent with those used in other clinical investigations [27]. An ordinal scale was devised and although not tested for reliability the questions had high face validity. These assessments were not associated with treatment times.

Secondary outcome measures using the same instruments as well as health care utilization data were collected to evaluate long-term changes (up to 1 year). Assessment of low back biomechanics was also completed. These results will be reported in separate papers.

Sample size

The sample size was based on a prior randomized clinical trial of chronic low back pain patients comparing manipulation to a sham treatment and back school [31]. Data from that investigation indicated that 200 patients would provide significance at $\alpha=0.05$ and $\text{power}=0.80$ for a change in perceived pain.

Randomization

Randomization occurred after the patient had been assessed for all inclusion/exclusion criteria. Randomization occurred before baseline outcome measures for the study were undertaken.

Sequence generation

Random number tables were used to generate treatment allocation.

Allocation concealment

Sequentially numbered sealed manila envelopes held each successive randomized treatment group allocation. At the time of randomization the research assistant opened the next numbered envelope and the subject was allocated accordingly.

Implementation

The allocation sequence was generated by the clinical coordinator. Patient enrollment was the responsibility of the research assistants at the chiropractic institution where patients had been evaluated by attending chiropractic clinicians. Neither the clinician who first saw the

patient nor the patient who agreed to participate in the study was involved in the allocation to intervention group. Research assistants were not blinded to patient allocation; however, as described below, they were blinded to the outcome measures, and the primary measures were self-administered questionnaires that were not completed in the presence of the attending clinician at any time.

Blinding

The primary outcome measures were self-administered questionnaires distributed by the research assistants. Study participants were given blank questionnaires at each assessment point and placed completed forms in an envelope. Subjects then sealed the envelope and returned it to the research assistant. Research assistants remained blinded to outcome data for the entire study period and were counseled by the research investigators and clinical coordinator, regarding the importance of blinding. They were trained in administration of informed consent and outcome data retrieval using simulated patients. Meetings between the research co-coordinator, principal investigator and providers responsible for treatment were held on a regular basis throughout the study to facilitate quality control. No incidents of unblinding were reported.

Statistical methods

The primary outcomes were evaluated by ANCOVA. Bonferroni correction was applied for repeated testing in the primary measures, and the alpha level was therefore set to 0.01. Secondary analysis was completed on questions related to satisfaction. Additional subgroup analysis was performed as part of feasibility issue development. Subgroup categorization was based on physical examination, history, and clinician's assessment of severity and recurrence of pain episodes. *t* tests and χ^2 supplemented descriptive analyses providing information about potential trends.

Results

Participant flow

Figure 3 is a diagram of stages of study participation. There were no crossovers. A total of 2,176 subjects were screened. Of these, 1,941 (89.2%) did not meet one or more inclusion criteria (see Table 2) and 77 (3.5%) were eligible but chose not to sign the informed consent. The rather high number of subjects who presented with no low back pain resulted from monitoring all new patients

at both the chiropractic and orthopedic clinics. Thus, even patients with a different chief complaint were screened. The remaining 235 subjects (10.8%) were randomized. Of these, 123 subjects were allocated to FD and 112 to ATEP.

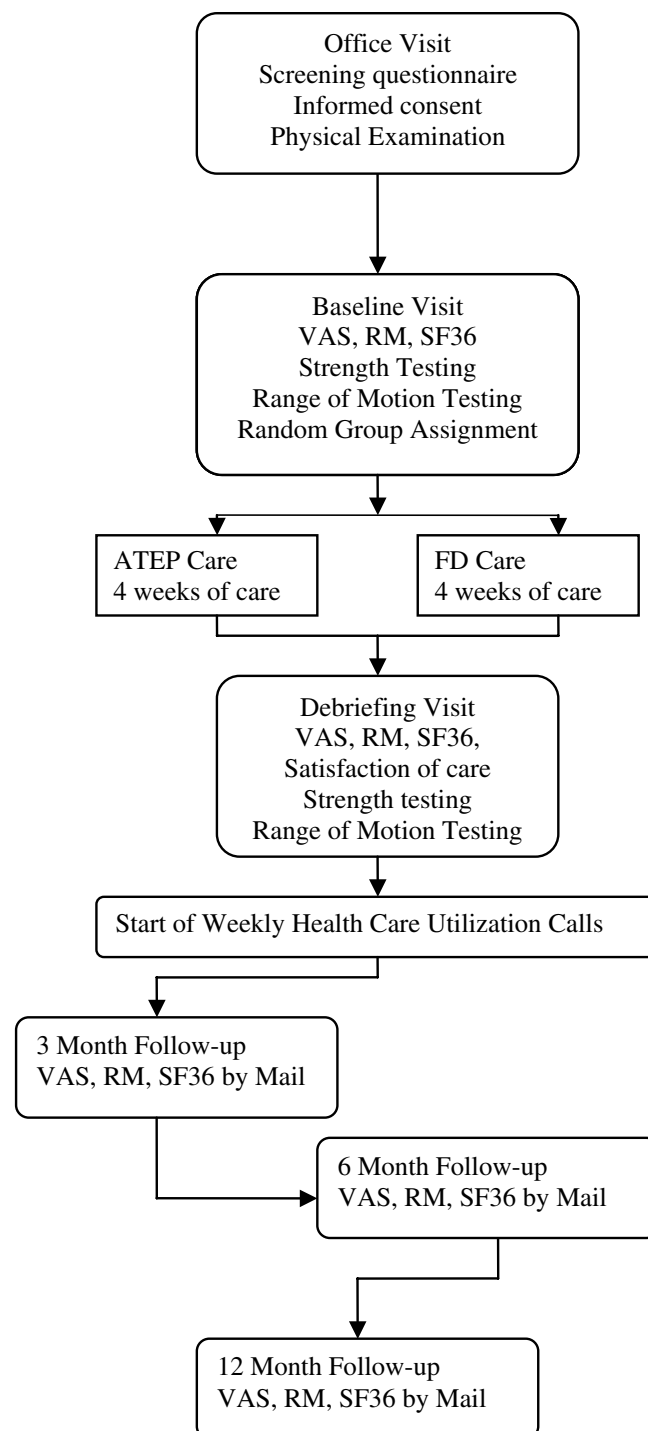


Fig. 3 Flow chart of study design

Table 3 Reasons for withdrawal

Reason	FD	ATEP
Refused care in randomized group	0	3
Scheduling conflicts	3	5
No longer interested in participation	7	14
Too much pain to continue	1	2
Injury or surgery unrelated to LBP	2	1
Total	13	25

A total of 197 subjects (83.4%) completed the intervention phase. Missing data (not completing or incorrectly completing the forms) resulted in 194 data points for the pre–post intervention VAS, 196 for the RM, and for each of the SF-36 subscales as follows: physical function 194; role function physical problems 192; bodily pain 193; general health 188; vitality fatigue 190; social functioning 194; role function emotional problems 192; mental health 190; physical component score 188; and mental component score 188.

Of the 38 dropouts, 13 were from FD and 25 from ATEP. Primary reasons for study withdrawal were diminished interest and scheduling difficulties. Table 3

provides these data according to group membership. A difference in proportions test indicated that significantly more subjects dropped out of the study from ATEP ($P < 0.02$). The majority listed “no longer interested in participation” as their reason for withdrawal. Thus, increasing pain and adverse reactions did not account for this observation.

Recruitment

Recruitment of study participants began in August 1998 and was completed in December 1999.

Baseline data

Table 4 lists the baseline characteristics of subjects by group. No significant differences were found, although a trend towards more individuals with lower household incomes in the ATEP group was seen. Table 5 provides the baseline characteristics of subjects according to history of low back pain and primary outcome measures.

Table 4 Baseline characteristics of eligible subjects according to treatment allocation

Variable	FD			ATEP			<i>P</i> value ^a
	Number of subjects	Mean or percent	SE	Number of subjects	Mean or percent	SE	
Age	123	42.22	1.03	112	40.88	1.21	0.40
Gender							
Males	81	65.85		66	58.93		0.27
Females	42	34.15		46	41.07		
Race							
Caucasian	102	82.93		91	81.98		0.44
Hispanic	5	4.07		7	6.31		
African American	8	6.50		7	6.31		
Asian	8	6.50		4	3.60		
Other	0	0.00		2	1.80		
Marital status							
Married	74	60.16		66	58.93		1.00
Divorced	20	16.26		19	16.96		
Widowed	3	2.44		3	2.68		
Never married	26	21.14		24	21.43		
Education							
< High school diploma	3	2.44		4	3.60		0.43
High school diploma	53	43.09		53	47.75		
Trade or technical school	18	14.63		16	14.41		
Bachelors degree	32	26.02		31	27.93		
MS/PhD/Other advanced degree	17	13.83		7	6.31		
Employment category							
Manual labor	42	34.15		32	28.57		0.60
Non-manual labor	66	53.66		63	56.25		
Unemployed, retired	15	12.20		17	15.18		
Total household income							
< \$20,000	6	4.96		16	15.09		0.07
\$20,000–39,999	28	23.14		25	23.58		
\$40,000–59,999	36	29.75		27	25.47		
≥\$60,000	51	42.15		38	35.85		

^aBased on *t* test or χ^2

Table 5 Baseline characteristics of eligible subjects related to low back pain

Variable	FD			ATEP			P value ^a
	Number of subjects	Mean or percent	SE	Number of subjects	Mean or percent	SE	
Onset							
Sudden	38	30.89		34	30.36		0.93
Gradual	85	69.11		78	69.64		
Number of prior episodes							
None (first episode)	29	23.58		17	15.18		0.13
One	10	8.13		4	3.57		
Two or three	20	16.26		19	16.96		
Four or more	64	52.03		72	64.29		
Disability							
On or applied for	0	0.00		4	0.04		0.05
Not on	112	100.00		102	0.96		
Workers' compensation							
On or applied for	2	0.02		5	0.05		0.27
Not on	110	0.98		101	0.95		
Unsettled litigation for low back pain							
Yes	1	0.01		1	0.01		1.00
No	118	0.99		110	0.99		
Radiculopathy							
Yes	22	17.89		23	20.54		0.61
No	101	82.11		89	79.46		
Zung depression score	123	43.34	0.81	112	45.06	0.89	0.15
VAS	123	38.00	2.01	112	35.70	1.96	0.42
Roland Morris	122	6.64	0.43	112	6.84	0.42	0.74
SF-36							
Physical function	122	69.87	1.76	112	69.83	2.11	0.99
Role function, physical problems	120	45.83	3.44	112	50.00	3.68	0.41
Bodily pain	121	47.36	1.42	112	44.96	1.48	0.24
General health	120	70.71	1.71	111	69.87	1.84	0.74
Vitality/fatigue	122	51.38	1.83	111	48.60	2.02	0.31
Social functioning	122	77.46	1.87	112	74.22	2.47	0.29
Role function, emotional problems	120	75.00	3.31	112	71.13	3.62	0.43
Mental health	122	75.89	1.33	111	71.21	1.76	0.03
Physical component score	120	41.77	0.74	111	42.71	0.84	0.40
Mental component score	120	51.18	0.83	111	48.49	1.19	0.06

^aBased on *t* test or χ^2

The variables in Tables 4 and 5 were evaluated both with and without withdrawals. As no significant differences existed regardless of whether withdrawing subjects were included, data are presented for all randomized subjects.

A trend toward more prior episodes may have existed for the ATEP group. The Mental Health Scale of the SF-36 suggested a possible difference between groups based on initial scores. Thus, ANCOVAs for the VAS and RM were completed with and without this variable as a covariate. Inclusion did not significantly contribute to the model nor affect the outcome. It was, therefore, not included in the models reported here.

Numbers analyzed

Analysis for primary outcome measures (VAS, RM, and SF-36 component scores) at the 4-week follow-up was completed both using an intention-to-treat approach

and analyzing only those subjects completing the intervention phase and providing appropriately completed outcome forms. This was done because there appeared to be more dropouts in the ATEP group. Using only an intention-to-treat approach could have resulted in a bias against the ATEP since more of the results in this group would have been carried forward in subjects who had not completed that therapy protocol.

Subjects provided data for the VAS on a weekly basis as part of their clinical protocol. Thus for the intention-to-treat analysis, the last available VAS score was used to calculate the results. For the RM and the SF-36 component scores, intention-to-treat analysis was completed by taking forward the outcome measure at baseline for all subjects whose data were missing. For the evaluation of those finishing the intervention phase, complete data were available for analysis at both the initial and the post 4-week intervention period. Numbers analyzed for the intention-to-treat evaluation and for

Table 6 Difference in pre- and post-intervention scores at 4 weeks, for primary outcome measures

Outcome	FD			ATEP			<i>P</i> value ^a	Effect size
	Number of subjects	Mean or percent	SE	Number of subjects	Mean or percent	SE		
VAS (intention to treat)	123	20.57	2.00	112	12.34	1.80	0.00	0.41
VAS (intervention completed)	108	22.66	2.12	86	15.46	2.20	0.01	0.35
Roland Morris (intention to treat)	123	2.81	0.38	112	2.30	0.33	0.17	0.11
Roland Morris (intervention completed)	109	3.17	0.41	87	2.97	0.39	0.68	0.04
SF-36 scales (intervention completed)								
Physical function scale	109	11.82	1.97	85	9.43	1.65		
Role function/physical probability scale	107	25.70	3.82	85	29.22	4.23		
Bodily pain scale	108	13.48	1.72	85	17.27	1.90		
General health scale	104	3.68	1.06	84	4.32	1.26		
Vitality/fatigue scale	106	10.16	1.34	84	8.45	1.51		
Social functioning scale	109	10.67	2.01	85	10.00	1.89		
Role function/emotional probability scale	107	11.84	3.46	85	10.20	4.06		
Mental health scale	106	4.44	1.00	84	4.88	1.51		
SF-36 component scores								
Physical component score (intention to treat)	120	5.02	0.66	111	4.78	0.62	0.89	0.03
Physical component score (intervention completed)	104	5.80	0.74	84	6.32	0.74	0.33	0.07
Mental component score (intention to treat)	120	2.11	0.62	111	1.69	0.71	0.12	0.06
Mental component score (intervention completed)	104	2.43	0.71	84	2.23	0.94	0.63	0.17

^aANCOVA

those completing the intervention phase are listed in Table 6.

Outcomes and estimation

Significant differences were observed in the pre to post measures for all primary outcomes at 4 weeks, regardless of treatment group (VAS: $t = 12.58$, $P < 0.01$; RM: $t = 10.73$, $P < 0.01$; SF-36 physical component score: $t = 11.50$, $P < 0.01$; SF-36 mental component score:

$t = 4.08$, $P < 0.01$). Table 6 provides the data regarding pre-post intervention difference scores related to the two groups with respect to the primary outcomes. ANCOVA, accounting for pre-treatment scores, indicated a statistically significant difference in the VAS between the two treatment groups ($F = 6.18$, $P = 0.01$), favoring FD. No such difference was observed for either the RM or the two component scores of the SF-36. Figure 4 provides box and whisker plot of the unpaired data for both the VAS and the RM at baseline and after intervention.

Fig. 4 Box and whisker plot of the VAS and Roland Morris scores at baseline and after intervention—unpaired data

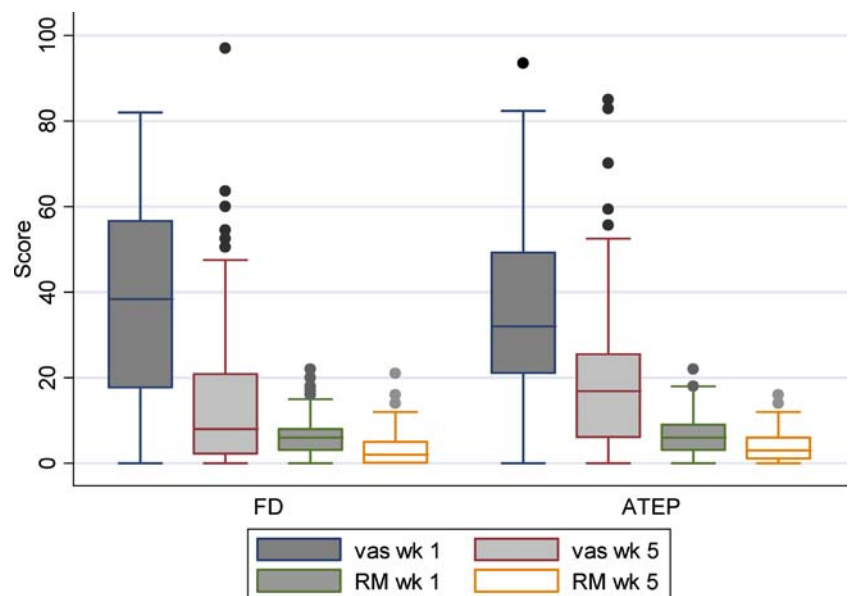


Table 7 Satisfaction

Variable	FD		ATEP		<i>P</i> value ^a
	Number of subjects	Percent	Number of subjects	Percent	
Overall how much were you helped?					
Not at all	3	2.91	2	2.41	0.11
A little bit	15	14.56	24	28.92	
Quite a bit	30	29.13	26	31.33	
Very much	52	50.49	28	33.73	
Not sure	3	2.91	3	3.61	
In the future, would you return to this type of care?					
Definitely	55	53.92	31	37.35	0.16
Probably	32	31.37	33	39.76	
Not sure	11	10.78	14	16.87	
Unlikely	4	3.92	5	6.02	
Would you recommend this type of care to family or friends?					
Definitely	62	60.19	37	44.57	0.15
Probably	28	27.18	27	32.53	
Uncertain	10	9.71	15	18.07	
Probably not	3	2.91	4	4.82	

^a χ^2

Ancillary analyses

Table 7 presents data regarding subject satisfaction according to treatment group. A large proportion of members in both groups felt they were helped and would be willing to return to the type of care they were allocated to. No statistical difference was observed between groups.

Table 8 provides the descriptive data for the VAS and RM at 3, 6, and 12 months. Although complete analysis of these data will be reported in a separate paper, initial gains made by both groups appear to remain relatively stable throughout the complete follow-up period. Although some subjects were lost to follow-up, data were retrieved from 78% (96/123) subjects in the FD group and 70% (78/112) of subjects in the ATEP group after 12 months. It is likely that the higher rate of follow-up at 12 months was achieved as a result of the monetary incentive subjects had for providing information at that time point.

In an effort to guide future multi-center trials in appropriately matching study patients to contrasting interventions, subgroup analysis was performed. Subjects were first divided according to the presence of radiculopathy. Results in Table 9 suggest a trend toward greater improvement in perceived pain for those with radiculopathy, in the FD group. Subjects were then divided according to the level of severity and according to whether the term “chronic” referred to continually present pain or recurrent pain. These results are presented in Table 10. The majority of subjects were defined as having continually present pain.

Descriptive results suggested that moderate to severely affected continual chronic pain patients may have benefited most from FD. When data from these moderate and severe subgroups were combined, a 27.22% change in VAS was observed in the FD group compared with only a 14.36% change in the ATEP group ($t = 3.09$, $P = 0.00$). Although a minority of patients was classified

Table 8 Follow-up descriptive data for VAS and Roland Morris (RM): change from baseline to time periods indicated

Variable	Group	Time								
		3 months			6 months			12 months		
		Number of subjects	Mean	SE	Number of subjects	Mean	SE	Number of subjects ^a	Mean	SE
VAS	FD	87	16.52	2.95	90	18.26	2.64	96	17.10	2.55
RM	FD	86	3.50	0.50	90	3.89	0.46	95	3.90	0.53
VAS	ATEP	76	12.04	2.53	74	8.92	2.89	78	12.36	2.43
RM	ATEP	76	3.75	0.51	78	3.42	0.50	78	3.77	0.44

Statistical analysis of long-term follow-up data is currently being undertaken in a separate paper

^aLarger number of subjects at 12-month follow-up coincides with monetary reward provided to subjects for study completion

Table 9 Change in perceived pain (VAS) for patients with and without radiculopathy according to intervention group

Radiculopathy	FD			ATEP			<i>P</i> value ^a
	Number of subjects	Mean	SE	Number of subjects	Mean	SE	
Yes	19	26.47	4.93	19	10.91	5.73	0.05
No	89	21.85	2.34	67	16.75	2.32	0.13

^a*t* tests

as having recurrent pain, the descriptive statistics in Table 10 suggests the opposite result for this subgroup. When data from the moderate and severe subgroups in this category were combined, a 12.99% change in VAS was observed in the FD group, while a 25.81% change was observed in the ATEP group. Although this result was not statistically significant ($t=1.51$, $P=0.15$), the effect size was substantial (0.87). Using this effect size, an alpha level of 0.01, and 80% power, a sample size of approximately 30 recurrent, moderate to severely affected subjects would be required for a future test.

Adverse events

No adverse events or side effects were reported from subjects in either intervention group.

Discussion

The first objective of this investigation was to compare the effectiveness of chiropractic FD to physical therapy ATEP by evaluating change in perceived pain, function, and health status before and after 4 weeks of intervention. Like the study of other treatments for chronic low back pain [17], immediate and significant relief of pain was considered the primary clinically relevant outcome. Results indicated statistically significant gains to both groups. In addition, subjects in the FD group were ob-

served to experience significantly greater reduced perceived pain (according to VAS measures) after treatment ($P=0.01$). Although function (via the Roland Morris) was not significantly different for this group compared with ATEP, musculoskeletal pain remains an important intangible cost associated with low back syndromes [19] and its reduction provides additional clinically relevant data for future consideration. One potential limitation to any study is the sample selection process. Although subjects were recruited using a variety of techniques and from a variety of environments, there is no way to know if this or any other sample is unique and thus biased. It may be, for example, that patients responding to the media advertisements or presenting to the chiropractic clinics were hoping to try one or the other form of care. Allocation to the opposite treatment group may not have met patient expectations and could bias against the treatment outcome for that group, while allocation to the patient-preferred group could result in a bias in the opposite direction.

Many forms of treatment are used in everyday practice for chronic low back pain. Reports of clinical trials using physical modalities such as manipulation [23] and exercise, however, often fail to provide necessary details allowing either reproduction in a clinical setting or comparison of trials to judge the preponderance of evidence for policy making. Heterogeneity of therapeutic procedures, for example, choices among exercise routines [32] or massage therapies [10], has hampered attempts at formal meta-analyses. This may, in part,

Table 10 Change in perceived pain (VAS) associated with low back pain subgroup and contrasting study interventions

Low back subgroup	FD			ATEP			<i>P</i> value ^a
	Number of subjects	Mean	SE	Number of subjects	Mean	SE	
All chronic	91	23.75	2.25	71	14.38	2.25	0.00
Mild	27	15.53	4.71	25	14.41	3.00	0.85
Moderate	59	27.33	2.36	42	14.68	3.34	0.00
Severe	5	25.87	14.30	4	11.03	6.23	0.42
Moderate and severe combined	64	27.22	2.40	46	14.36	3.09	0.00
All recurrent	17	16.85	5.95	14	18.89	7.13	0.83
Mild	3	34.83	26.38	4	1.60	14.62	0.29
Moderate	12	16.08	4.73	8	21.94	8.76	0.53
Severe	2	-5.51	16.22	2	41.31	7.65	0.12
Moderate and severe combined	14	12.99	4.85	10	25.81	7.47	0.15

^a*t* tests

explain conflicting results among meta-analyses for various treatment strategies [2, 10, 22, 23, 25, 26]. Our clinical trial compared the use of a flexion–distraction intervention with a specific active trunk exercise routine (ATEP). Details regarding these treatments are provided for future consideration of these data relative to additional studies.

While disparity within treatment modalities, between clinical trials of chronic low back pain may have facilitated conflicting evidence, additional problems can be attributed to substantial variation in the definition of the word “chronic”, as related to low back pain. Andersson [1], for example, found five distinct definitions:

1. Pain lasting longer than 7–12 weeks
2. Pain lasting longer than expected
3. Recurrent back pain
4. Symptoms resulting in loss of days at work or disability
5. Convenience diagnosis for individuals disabled for other reasons (including psychological)

Thus, both time frame and “recurrence” vary among studies. Further complicating this issue is symptom severity. Subjects with mild pain may respond differently to a single treatment strategy in comparison to subjects who are in severe pain. Studies of patients with chronic low back pain may or may not include patients with leg symptoms, and these may or may not be related to frank radiculopathy. Chronic low back pain then is an extremely complex syndrome.

It has been suggested that patients be classified into subgroups [4, 18, 33] and that inclusion into clinical trials consider treatment according to the prognosis of individual subgroups in specific treatment regimes [4, 9].

In response to the above concerns, the second objective of this study was to consider the data from the perspective of a feasibility study designed to generate hypotheses. The multidisciplinary team of chiropractors, allopathic physicians, and physical therapists assembled succeeded in its collaborative efforts to implement protocols and complete the study as designed. Sample diversity allowed evaluation of treatment effects for a variety of subcategories generated by assessing characteristics of pain such as radiculopathy, chronicity, and severity. When study subjects were divided according to the presence of radiculopathy, results suggested some benefit for patients with radiculopathy who received FD. Further, subjects categorized as “chronic” with moderate to severe symptoms appeared significantly more improved post-FD than post-ATEP ($P < 0.01$). Interestingly, however, if subjects were categorized as “recurrent” with moderate to severe symptoms, descriptive changes suggested greater improvement of almost exactly the same magnitude but favoring ATEP. Small sample sizes are, of course, a limitation to reach-

ing any conclusion from subgroup analysis. Rather, these data should be used to provide information to guide future study.

The differences in treatment results according to subgroup analyses make biological sense. The FD intervention was intended to provide motion and forces directed at specific intervertebral levels [14]. The ATEP on the other hand was intended to concentrate more on strengthening the muscles surrounding the spine and increasing flexibility. As such, a greater decrease in VAS among patients with radiculopathy should be expected for the FD group where changes in disc pressure [7, 12, 13] may be most important. The diametrically opposed results dependent on patient categorization as moderate to severely chronic versus moderate to severely recurrent provide interesting data for consideration. Clinical trial outcomes may vary, according to the representation of previously undefined subgroups. In randomized clinical trials where balance is expected due to the randomization procedure, the mean change within each intervention will tend towards zero if subgroups have opposite responses. Slight changes in this balance in either direction would produce conflicting results, consistent with conclusions from systematic review [18]. Since no single investigation such as this can be expected to define the nature of low back subgroups, future extensive epidemiologic study of this area is suggested.

Conclusion

In accordance with many studies of chronic low back pain, patients perceived significantly less pain after intervention, regardless of group allocation. Subjects randomly allocated to FD had significantly greater relief from perceived pain, as defined by VAS scores, than those in ATEP. According to the Roland Morris, both groups responded in terms of function and there was no difference between groups on this measure. Subgroup analysis indicated that subjects categorized as chronic, with moderate to severe symptoms, and those with radiculopathy improved most with FD. Subjects categorized with recurrent pain and moderate to severe symptoms improved most with ATEP. This may help explain contrasting outcomes among previous trials of chronic low back pain treatments.

Acknowledgements The authors thank the Health Resources and Services Administration (HRSA) for their financial support (Grant # R18 AH 10001), National Chiropractic Mutual Insurance Company, and many chiropractic physicians for their generous donations. We also thank James Cox, DC and Timothy Carey, PhD, Rita Ator, PT for providing invaluable consultation and the clinicians, physical therapists, student assistants, study patients, and clinic support staff for their help with this study.

References

1. Andersson GB (1999) Epidemiological features of chronic low-back pain. *Lancet* 354:581–585
2. Assendelft WJ, Morton SC, Yu EI, Suttorp MJ, Shekelle PG (2003) Spinal manipulation for low back pain. A meta-analysis of effectiveness relative to other therapies. *Ann Intern Med* 138:881
3. Beaton DE, Hogg-Johnson S, Bombardier C (1997) Evaluating changes in health status: reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. *J Clin Epidemiol* 50:79–93
4. Childs JD, Fritz JM, Flynn TW, Irrgang JJ, Johnson KK, Majkowski GR, Delitto A (2004) A clinical prediction rule to identify patients with low back pain most likely to benefit from spinal manipulation: a validation study. *Ann Intern Med* 141:920–928
5. Christensen MG, Kerkoff D, Kollasch MW (2000) Job analysis of chiropractic, 2000. National Board of Chiropractic Examiners, Greeley
6. Cox JM (1999) Low back pain: mechanism, diagnosis, treatment. 6th edn. Lippincott Williams & Wilkins, Baltimore
7. Cox JM, Gudavalli MR (2005) Traction and distraction techniques. In: Haldeman S, Dagenais S, Budgell B, Grunnet-Nilsson N, Hooper PD, Meeker WC, Triano J (eds) Principles and practice of chiropractic. 3rd edn. McGraw-Hill, New York, pp 821–840
8. Deyo RA, Phillips WR (1996) Low back pain. A primary challenge. *Spine* 21:2826–2832
9. Flynn T, Fritz J, Witman J, Wainner R, Magel J, Rendeiro D, Butler B, Garber M, Allison S (2002) A clinical prediction rule for classifying patients with low back pain who demonstrate short-term improvement with spinal manipulation. *Spine* 27:2835–2843
10. Furlan AD, Brosseau L, Imamura M, Irvin E (2002) Massage for low back pain. *Cochrane Database Syst Rev* (2):CD001929
11. Gatchel RJ (ed) (2001) Compendium of outcome instruments for assessment and research of spinal disorders. North American Spine Society, LaGrange
12. Gudavalli MR, Cox JM, Baker JA, Cramer GD, Patwardhan AG (1997) Intervertebral disc pressure changes during the flexion–distraction procedure for low back pain. In: Proceedings of the 1997 annual International Society for the Study of the Lumbar Spine, Singapore, p 165
13. Gudavalli MR, Cox JM, Baker JA, Cramer GD, Patwardhan AG (1997) Intervertebral disc pressure changes during a chiropractic procedure for low back pain. American Society of Mechanical Engineers bioengineering conference, Dallas, TX, pp 215–216
14. Gudavalli MR, Cox JM, Cramer GD, Baker JA, Patwardhan AG (2000) Vertebral motions during flexion–distraction treatment for low back pain. 2000 ASME international mechanical engineering congress and exposition, Orlando, FL, pp 129–130
15. Helmhout PH, Harts CC, Staal JB, Candel MJ, de Bie RA (2004) Comparison of a high-intensity and a low-intensity lumbar extensor training program as minimal intervention treatment in low back pain: a randomized trial. *Eur Spine J* 13:537–547
16. Huskisson EC (1982) Measurement of pain. *J Rheumatol* 9:768–769
17. Katz N, Rodgers DB, Krupa D, Reicin A (2004) Onset of pain relief with rofecoxib in chronic low back pain: results of two four-week, randomized, placebo-controlled trials. *Curr Med Res Opin* 20:651–658
18. Koes BW, Assendelft WJ, van der Heijden GJ, Bouter LM (1996) Spinal manipulation for low back pain. An updated systematic review of randomized clinical trials. *Spine* 21:2872–2873
19. Lubeck DP (2003) The costs of musculoskeletal disease: health needs assessment and health economics. *Best Pract Res Clin Rheumatol* 17:529–539
20. Mayer TG, Gatchel RJ, Evans TH (2002) Chronic low back pain. In: Fitzgerald RH, Kaufer H, Malkani AL (eds) Orthopaedics. Mosby, St Louis, pp 1192–1197
21. McDowell I, Newell C (1996) Measuring health: a guide to rating scales and questionnaires. Oxford University Press, New York
22. Mior S (2001) Exercise in the treatment of chronic pain. *Clin J Pain* 17:S77–S85
23. Mior S (2001) Manipulation and mobilization in the treatment of chronic pain. *Clin J Pain* 17:S70–S76
24. Murphy DR, Morris C (2005) Manual examination of the patient. In: Haldeman S, Dagenais S, Budgell B, Grunnet-Nilsson N, Hooper PD, Meeker WC, Triano J (eds) Principles and practice of chiropractic. 3rd edn. McGraw-Hill, New York, pp 593–610
25. Nielson WR, Weir R (2001) Biopsychosocial approaches to the treatment of chronic pain. *Clin J Pain* 17:S114–S127
26. Niemisto L, Kalso E, Malmivaara A, Seitsalo S, Hurri H (2003) Radiofrequency denervation for neck and back pain. A systematic review of randomized controlled trials. *Cochrane Database Syst Rev* (1):CD004058
27. Ohnmeiss DD, Rashbaum RF (2001) Patient satisfaction with spinal cord stimulation for predominant complaints of chronic, intractable low back pain. *Spine J* 1:358–363
28. O’Sullivan PB (2000) Lumbar segmental ‘instability’ clinical presentation and specific stabilizing exercise management. *Man Ther* 5:2–12
29. O’Sullivan PB, Phytz GD, Twomey LT, Allison GT (1997) Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylolysis or spondylolisthesis. *Spine* 22:2959–2967
30. Roland M, Morris R (1983) A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability and low back pain. *Spine* 8:141–144
31. Triano JJ, McGregor M, Hondras MA, Brennan PC (1995) Manipulative therapy versus education programs in chronic low back pain. *Spine* 20:948–955
32. van Tulder M, Malmivaara A, Esmail R, Koes B (2000) Exercise therapy for low back pain: a systematic review within the framework of the Cochrane collaboration back review group. *Spine* 25:2784–2796
33. Turk DC (2005) The potential of treatment matching for subgroups of patients with chronic pain: lumping versus splitting. *Clin J Pain* 21:44–55
34. Vollenbroek-Hutton MMR, Hermens HJ, Wever D, Gorter M, Rinket J, Ijzerman MJ (2004) Differences in outcome of a multidisciplinary treatment between subgroups of chronic low back pain patients defined using two multi-axial assessment instruments: the multidimensional pain inventory and lumbar dynamometry. *Clin Rehabil* 18:566–579