SPINE Volume 29, Number 23, pp 2593–2602

Does it Matter Which Exercise?

A Randomized Control Trial of Exercise for Low Back Pain

Audrey Long, BScPT,* Ron Donelson, MD,† Tak Fung, PhD‡

Study Design. Multicentered randomized controlled trial.

Objectives. To determine if previously validated low back pain (LBP) subgroups respond differently to contrasting exercise prescriptions.

Summary of Background Data. The role of "patient-specific" exercises in managing LBP is controversial.

Methods. A total of 312 acute, subacute, and chronic patients, including LBP-only and sciatica, underwent a standardized mechanical assessment classifying them by their pain response, specifically eliciting either a "directional preference" (DP) (*i.e.*, an immediate, lasting improvement in pain from performing either repeated lumbar flexion, extension, or sideglide/rotation tests), or no DP. Only DP subjects were randomized to: 1) directional exercises "matching" their preferred direction (DP), 2) exercises directionally "opposite" their DP, or 3) "nondirectional" exercises. Outcome measures included pain intensity, location, disability, medication use, degree of recovery, depression, and work interference.

Results. A DP was elicited in 74% (230) of subjects. One third of both the opposite and non-directionally treated subjects withdrew within 2 weeks because of no improvement or worsening (no matched subject withdrew). Significantly greater improvements occurred in matched subjects compared with both other treatment groups in every outcome (P values <0.001), including a threefold decrease in medication use.

Conclusions. Consistent with prior evidence, a standardized mechanical assessment identified a large subgroup of LBP patients with a DP. Regardless of subjects' direction of preference, the response to contrasting exercise prescriptions was significantly different: exercises matching subjects' DP significantly and rapidly decreased pain and medication use and improved in all other outcomes. If repeatable, such subgroup validation has important implications for LBP management.

Key words: exercise, centralization, low back pain subgroups, directional preference, McKenzie method. **Spine 2004;29:2593–2602**

A number of systematic reviews have raised important questions regarding the role of exercise in the treatment of low back pain (LBP) with a lack of evidence supporting any specific type of exercise, *e.g.*, back or abdominal strengthening, McKenzie, Williams, flexion, extension, or stretching.[1–5] LBP clinical guidelines advocate advice to stay active and an early return to normal activity as the means to faster recovery with less disability.[6–8] These guidelines challenge the popular clinical practice of prescribing patient-specific exercises[9,10] determined by an individual's assessment findings and implies that nonspecific exercise can be prescribed without consideration of individual clinical signs, *e.g.*, every LBP patient is given the same exercises.

There is growing opinion that the equivocal or conflicting results among exercise trials can be attributed to the faulty assumption that the "nonspecific" LBP populations studied were homogeneous.[6,11–14] Alternatively, treatment efficacy has been demonstrated in trials that studied defined LBP subgroups, although methodologic weaknesses require cautious interpretation of results.[15–20] Meanwhile, identification of LBP subgroups was listed as the top research priority by the International Forum Primary Care Research in Low Back Pain.[21] The Cochrane Back Review Group has recently referred to the identification of subgroups and predictors of chronicity as "the Holy Grail" of LBP.[22]

One subgroup classification method (McKenzie Method),[23] often referred to as Mechanical Diagnosis and Therapy (MDT), has demonstrated strong inter-rater reliability (kappa values ranging from 0.79 to 1.0)[17,24–30] and other clinically useful properties: predicting outcome and discogenic pathology, and providing preliminary evidence of patient-specific treatments based on assessment findings.[15,24,29,31–35]

An important feature of the MDT assessment is the identification of a patient's "directional preference" (DP).[23,36–39] DP is identified when posture or repeated end-range movements in a single direction (flexion, extension, or side-glide/rotation) decrease or abolish lumbar midline pain, or cause referred pain emanating from the spine to appear to progressively retreat in a proximal direction back toward the lumbar midline ("centralization"). There is often a rapid and concurrent restoration of lumbar range of movement.[23] The inter-rater reliability for identifying DP in the hands of qualified practitioners (McKenzie Institute credentialed) is reported as excellent (agreement 90%, kappa 0.9).[26]

The objective of this study is to determine if a subject-specific exercise prescription concordant with a study participant's DP will achieve better outcomes than non-concordant exercises.

From *Bonavista Physical Therapy, Calgary, Alberta, Canada; †Self-Care First, Hanover, NH; and ‡Department of Psychology, University of Calgary, Calgary, Alberta, Canada.
Supported by the Community Ethics Review Board of the Alberta Heritage Foundation for Medical Research, the Physiotherapy Foundation of Canada, McKenzie Institute International, and Cambridge Physiotherapy Associates.
The manuscript submitted does not contain information about medical device(s)/drug(s).
Foundation funds were received to support this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.
Address correspondence and reprint requests to Audrey Long, BScPT, 620 Willesden Dr. S.E., Calgary, Alberta, Canada, T2J 2G1; E-mail: longma@telusplanet.net

Table 1. Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
Consecutive patients with: Low back pain With or without leg symptoms With or without one neurological sign Age 18–65 years Demonstrating a directional preference during the mechanical assessment	Cauda equina 2 or more neurological signs Spinal fractures Post-surgical Off work for 1 year or more due to LBP Medical causes (*e.g.,* severe osteoporosis, inflammatory or infectious conditions) Uncontrolled medical conditions (*e.g.,* diabetes, angina, hypertension) Pregnancy Inability to read English* Patients with prior knowledge of, or specific physician referral for, the McKenzie method were excluded to control for patient bias No directional preference elicited

*Study participants from Germany were provided with translated questionnaires

■ Methods

Physical therapists (PTs) in 11 clinics volunteered to participate in response to advertisements placed in McKenzie Institute Branch newsletters. The Community Ethics Review Board of the Alberta Heritage Foundation for Medical Research granted approval for the study.

Consecutive patients presenting for treatment of their LBP were asked to participate in a study of exercise for LBP. The consent form stated: "previous studies have shown that exercise helps promote healing, can help control pain, and improve function. However, there are a variety of opinions regarding which exercise(s) is the best."

Baseline subject characteristics recorded included age, gender, marital status, prior episode history, job demands, current work status, mechanism of injury, current episode duration, symptom location, and back and leg bothersome ratings.[40]

Our inclusion/exclusion criteria are listed in Table 1. The qualifying process also included a standardized MDT assessment by credentialed or diplomaed McKenzie Institute-trained therapists to identify the subgroup of subjects who demonstrated a DP and to assign one of three directional labels to each individual's DP subset: extension DP (with or without hips positioned off-center), flexion DP, or a lateral DP (left or right side-glide or rotation in flexion).[23] Study participants were successfully shielded from recognizing their own directional preference by not discussing this concept with them during their baseline assessment. Their own directional label and the relevance of changes in distal pain/symptom location (centralization/peripheralization) were also not discussed. When pain location changes do occur, without specific education, patients routinely do not recognize it as beneficial, nor do they recognize the directional theme so commonly present.

After giving informed consent, each member of these three DP subsets (extension, flexion, and lateral) was randomized to one of three treatments using treatment allocation cards drawn by nonmedical staff from DP-labeled envelops. This ensured equal representation of each directional subset in each treatment group (Figure 1).

Treatments. The three treatment protocols, described in detail in Appendix A, are summarized as follows: 1) Matched: Subjects were taught unidirectional end-range lumbar exercises matching the direction of their DP identified during baseline assessment.[23] 2) Opposite: Subjects were also taught unidirec-

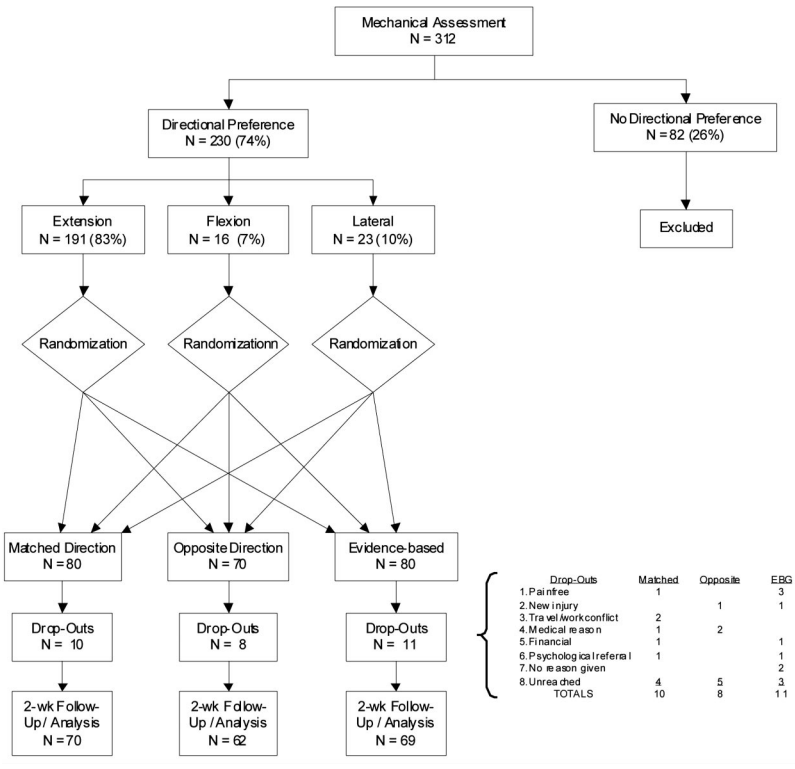

Figure 1. Study flow.

Table 2. Origin of Subjects By Country and Clinician.

Country	No. of Practices (all outpatient departments, except where stated otherwise)	No. of Patients	No. of Therapists by Practice (% by country)
Canada	2 Private practices	2	25/49 (24%)
United States	5	6	10/10/20/9/20 (22%)
Germany	1	1	88 (28%)
United Kingdom	2	2	14/61 (24%)
Kuwait	1	1	6 (2%)
Total	11	12	312 (100%)

tional end-range exercises, but in a direction opposite to their DP identified during baseline assessment. 3) Evidence-based care (EBC): Subjects were taught commonly prescribed multidirectional, midrange lumbar exercises, and stretches for the hip and thigh muscles.

Both the opposite and EBC groups were provided education consistent with LBP clinical guidelines, including advice aimed at minimizing fear avoidance behavior[41] and to remain active[8] (Appendix B). Members of the matched group were likewise instructed to remain active but also to avoid activities and positions that increase intensity or radiation of symptoms.

The baseline assessment in our study design required skills (*i.e.*, MDT) not part of standard PT training. The original design also called for the opposite and EBC treatments to be provided by non-McKenzie trained PTs. However, 6 (30%) of the first 20 subjects in our pilot project dropped out, 4 (67%) because of dissatisfaction with having to change therapists after the initial assessment. This suggested either success by the assessing therapists in convincing subjects there was no proven difference between the three forms of exercises, or some attractive rapport established during the assessment procedures (30–45 minutes). Since all MDT-trained therapists were also adequately trained to administer the other two exercise protocols, and since all had expectations that subjects could improve with all treatments (based at least on natural history and regression to the mean), subjects were then allowed to continue treatment with their assessing PT. Only 2 (9%) of the next 22 subjects dropped out. Furthermore, 2 years of clinician recruiting failed to identify sufficient clinics with both MDT and non-MDT trained clinicians working in the same clinic. Subsequently, all treating PTs were instructed to confidently present all three exercise programs, encouraging subjects' participation, with the intent of minimizing dropouts so as not to jeopardize the study. Data from these 42 pilot subjects were included in our analysis.

Treatment Protocols. For consistency, a minimum of three and maximum of six visits over the 2-week study period was imposed for all three treatment interventions. As in other studies,[29,31,32,35] our 2-week intervention period was deemed adequate time to determine how these three interventions addressed pain control. We also anticipated difficulty maintaining treatment compliance beyond 2 weeks (six visits) for those not experiencing improvement with their assigned treatment. Additionally, some clinics had contractual agreements to progress clients into progressive strengthening/functional programs. As part of their consent, subjects agreed not to participate in other nonmedicinal treatments during the trial and were advised that they were free to withdraw at any time. No attempts were made to influence medication use.

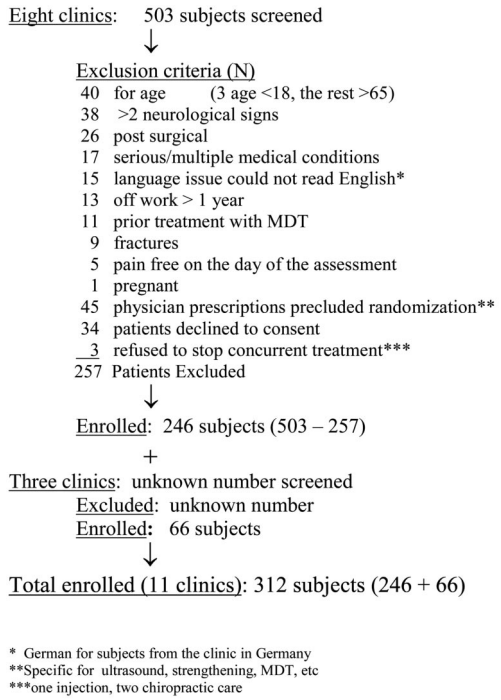

Figure 2. Exclusions.

Safety Guidelines. On recommendation of the Ethics Board, the participation of subjects reporting a lasting increase in pain,

Table 3. Subject Baseline Characteristics (N = 312)

Characteristic	No. (%)
Gender	
Male	166 (53)
Female	146 (47)
Prior episodes	
None	93 (30.0)
1 or 2	62 (19.9)
3 to 6	27 (11.9)
Many times, recover well in-between episodes	75 (24.0)
Many times, do not recover well between	31 (09.9)
Information not provided	24 (10.9)
Reported cause of injury*	
No known reason	180 (34.6)
Injury at home	74 (23.7)
Injury at work	74 (23.7)
Injury during sports	24 (7.7)
Car accident	10 (3.2)
Lift/twist	74 (23.7)
Push/pull	13 (4.2)
Fall	8 (2.6)
Major job demands*	
Student	5 (1.6)
Homemaker	30 (9.6)
Retired	14 (4.5)
Sitting	129 (41.3)
Light (<5 kg)	32 (10.30)
Medium (<23 kg)	43 (13.8)
Heavy (<50 kg)	49 (15.7)
Very heavy (>50 kg)	45 (14.4)

*Some study participants checked more than one category (*e.g.*, truck driver: sitting/heavy lifting).

Table 4. Baseline Subject Characteristics by Treatment Group

	Matched (N = 80) (%)	Opposite (N = 69) (%)	EBC* (N = 80) (%)	χ^2 (*df*)	*P*
Categorical baseline variable					
Gender					
Male	39 (48.8)	35 (50.7)	46 (57.5)	1.339 (2)	0.5119
Female	41 (51.3)	34 (49.3)	34 (42.5)		
QTF classification					
1	39 (48.8)	34 (49.3)	42 (52.2)	2.660 (6)	0.8501
2	14 (17.5)	14 (20.3)	17 (21.3)		
3	12 (15.0)	10 (14.5)	8 (10.0)		
4†	15 (18.8)	11 (15.9)	8 (10.0)		
QTF acute < 7 days	11 (14.1)	8 (11.4)	11 (14.7)	4.739 (4)	0.3151
QTF subacute < 7 weeks	31 (39.7)	18 (25.7)	22 (29.3)		
QTF chronic > 7 weeks	36 (46.2)	44 (62.9)	42 (56.0)	4.739 (4)	0.3151
Episodes					
History of prior episodes	59 (74.7)	48 (70.6)	55 (68.8)	0.7135	0.6999
This is a first episode	20 (25.3)	20 (29.4)	25 (31.3)		
Off work					
Yes	26 (36.6)	29 (44.6)	35 (46.7)	1.653	0.4376
No	45 (25.3)	36 (55.4)	40 (53.3)		
Taking medication					
Yes	28 (36.8)	33 (48.5)	32 (45.1)	2.139 (2)	0.3431
No	48 (63.2)	35 (51.5)	39 (54.9)		
Rating of pain interference with usual work (outside home and housework)					
1	4 (5.1)	1 (1.4)	1 (1.3)	8.9168 (8)	0.3494
2	10 (12.8)	13 (18.8)	12 (15.0)		
3	21 (26.9)	18 (26.1)	28 (35.0)		
4	30 (38.5)	23 (33.3)	32 (40.0)		
5	13 (16.7)	14 (20.3)	7 (8.8)		
Bothersome scale					
Back pain					
1	2 (2.5)	1 (1.4)	1 (1.3)	3.4292 (8)	0.9046
2	10 (12.5)	5 (7.2)	10 (12.5)		
3	22 (27.5)	24 (34.8)	34 (42.5)		
4	32 (40.0)	31 (44.9)	34 (42.5)		
5	14 (17.5)	8 (11.6)	11 (13.8)		
Leg pain					
1	25 (32.5)	22 (33.3)	29 (37.2)	2.052 (8)	0.9794
2	14 (18.2)	14 (21.2)	15 (19.2)		
3	21 (27.3)	15 (22.7)	15 (19.2)		
4	11 (14.3)	11 (16.2)	14 (17.9)		
5	6 (7.8)	4 (6.1)	5 (6.4)		
	Mean (SD)	Mean (SD)	Mean (SD)		
Continuous baseline variables					
Age	42.86 (9.55)	42.19 (10.34)	41.51 (10.76)	0.3454 (2,224)	0.7083
Length of current episode‡ (weeks)	13.70 (19.84)	17.65 (21.82)	14.55 (17.60)	0.7181 (2,199)	0.4890
Back pain rating/10	6.01 (2.36)	6.06 (2.24)	6.01 (2.05)	0.0101 (2,226)	0.9900
Leg pain rating§/10	4.59 (N=41) (2.50)	4.74 (N=42) (2.48)	4.78 (N=41) (2.56)	0.07 (2,121)	0.934
Medication Total pills per day ‖	3.37 (N=27) (2.92)	3.65 (N=17) (2.74)	3.10 (N=19) (2.64)	0.17 (2,60)	0.845
BDI-II	10.16 (9.03)	9.88 (9.35)	8.05 (6.92)	1.4355 (2,225)	0.2402

*Evidence-based care group.
†These QTF 4 subjects represent a 'skewed' QTF 4 sample. Subjects with > 1 sign were excluded.
‡A total of 22 subjects reported > 100 week-long episode (SE skew, 0.163). These subjects were excluded from the above data.
§Note floor effect: those with leg pain = zero at intake are excluded from this analysis.
‖Note floor effect: those not taking medication at intake are excluded from this analysis.

symptoms radiating more distally, or demonstrating any deterioration in neurologic signs/symptoms, would be discontinued with outcome measures obtained and their data included in our intention-to-treat analysis.

Therapists and Treatment Sites. Twelve PTs had a mean of 12 years experience (range, 3–18 years). All were credentialed or diplomaed in MDT for a mean of 3 years (range, 1–8 years), having passed at least the standardized, validated examination from the McKenzie Institute International, a level of MDT training shown to produce excellent reliability in patient assessment.[26,27,29,42] Each therapist's country of origin and study subject contribution are listed in Table 2. Contributions were skewed because of referral patterns, time committed to the project, job changes, and a maternity leave.

Outcome Measures. Primary outcomes included back and leg pain intensity ratings using an 11-point visual analogue scale,[43] the 24-item Roland Morris Disability Questionnaire,[44] and medication use, classifying subjects first as takers or nontakers of LBP medication. If taking medications, the total number of pills taken daily was recorded.

Secondary outcome measures included a rating of activity interference at work and home (0–5),[40] the 21-item Beck De-

Table 5. Characteristics of Dropouts Compared With Those Who Completed Treatment

Baseline Characteristic	Dropouts	Completed Trial	F or χ^2	P
Age (years)	41.1	42.3	T(226) = 0.59	0.5530
Gender				
Male	10 (34.5%)	156 (55.1%)	χ^2 (1) = 3.71	0.05*
Female	19 (65.5%)	127 (44.9%)		
Directional preference label				
Extension	23 (79.3%)	168 (83.6%)	χ^2 (2) = .5333	0.7659
Flexion	2 (6.9%)	14 (7.0%)		
Lateral	4 (13.8%)	19 (9.5%)		
Treatment group allocation				
Matched	10 (34.5%)	70 (34.8%)	F(2) = 0.184	0.9122
Opposite	8 (27.6%)	62 (30.8%)		
EBG	11 (37.9%)	69 (34.3%)		
QTF classification				
1	16 (55.2%)	100 (49.8%)	F(1) = 2.53	0.4695
2	5 (17.2%)	40 (19.9%)		
3	2 (6.9%)	33 (16.4%)		
4	6 (20.7%)	28 (13.9%)		
Mean length of current episode				
Acute	2 (6.9%)	30 (13.0%)	χ^2 (4) = 4.009	0.4047
Subacute	9 (31.0%)	71 (30.9%)		
Chronic	17 (58.6%)	122 (53.0%)		
Insufficient data	1 (3.5%)	7 (3.0%)		
Off work due to LBP	11 (44.0%)	79 (42.2%)	F(1) = .0278	0.8677
LBP intensity rating	6.55	5.95	T(228) = 1.37	0.1710
Leg pain intensity rating	4.82	4.72	T(140) = 0.16	0.8760
RMDQ–Likert	74.00	66.60	T(228) = 1.17	0.2450
RMDQ–traditional	18.35	17.75	T(228) = 0.53	0.5940
BDI-II	14.86	8.57	T(31.5) = 3.08	0.004*
Prior episodes = yes	15 (53.6%)	148 (74.0%)	χ^2 (1) = 4.08	0.043*

*Significant at $P = 0.05$.

pression Inventory (BDI),[4,5] and the Quebec Task Force severity rating (QTF 1–4).[3] The latter classifies patients by their pain location and neurologic status. A newly generated satisfaction questionnaire also rated subjects' response to treatment, readiness for discharge, need for further treatment, and the ability to return to work and leisure activities.

Outcome measures were administered at baseline and 2 weeks by nonmedical reception staff blinded to the study design. Outcome measures requiring direct therapist measurement (*e.g.*, range of movement, strength, straight-leg raise) were intentionally not used to minimize testing bias. However, when no nonmedical staff was available, the treating PT sometimes handed questionnaires to the subjects.

A take-home compliance-tracking sheet enabled subjects to record the number of exercise sessions per day. Compliance scores were rated as follows: 0 = poor; 1–2 sessions = fair; 3–4 sessions = good; 5+ = excellent.

Data Analysis. All study sites mailed their subject data sheets and questionnaires to our Alberta study center where a blinded technician manually entered the data. Random, double-entry methods documented entry accuracy.

Based on the effect sizes calculated, a minimal sample size of 38 per treatment group was required to have a power of 0.90 with an alpha level of 0.05 for hypothesis testing. A sample size of 300 was chosen to allow for dropout and subgroup analysis. Subjects unable/unwilling to continue for the full 2-week protocol were included in the analysis (intention to treat).

Descriptive statistics and frequencies distributions of all variables were determined. Specifically, two-way analysis of variance for all continuous dependent variables determined whether there was a 1) time by exercise group interaction effect, 2) time effect, and 3) exercise group effect. Simple effects testing was performed where appropriate. χ^2 tests determined relationships between groups and the categorical/ordinal dependent variables at 2 weeks. Correlation analyses determined the association among the interval/ratio variables. McNemar χ^2 was used to confirm the heterogeneity of the changes within the treatment groups.

■ Results

Eight of the 11 clinics evaluated 503 consecutive LBP patients with 257 excluded (Figure 2). The remaining three clinics contributed 64 subjects, however, provided no exclusion data. Baseline characteristics of the 312 recruited subjects are listed in Table 3. A DP was elicited in 230 (74%), with the remaining 82 (26%) excluded (Figure 1). After randomization, there were no differences among the three treatment groups in any baseline demographic characteristics or outcome measures (Table 4).

Of the 230 DP subjects, three directions of DP were identified during the baseline assessment: 191 (83%) extension, 16 (7%) flexion, and 23 (10%) lateral responders (Figure 1). Twenty-four of those labeled as having an extension DP first needed their pelvis offset from the midline before their prone extension exercises produced centralization or reduction of pain. Figure 1 illustrates the randomization process.

Dropouts

Twenty-nine subjects (12.6%) failed to return for treatment appointments and did not provide data at 2 weeks, equally distributed between the treatment groups. None

Table 6. Outcomes by Treatment Groups (ANOVA)

Outcome Measure	Pre/Post	Matched [mean (SD)]	Opposite [mean (SD)]	EBC* [mean (SD)]	Time Effect *F* (*df*) [*P*]	Group Effect *F* (*df*) [*P*]	Interaction *F* (*df*) [*P*]
Back pain/10	Pre	5.86	6.08	5.97	138.50 (1,200) [<0.001†]	8.09 (2,200) [<0.001†]	11.55 (2,200) [<0.001†]
	Post	(2.39)	(2.17)	(2.06)			
		2.51	4.65	4.34			
		(1.96)	(2.33)	(2.51)			
Leg pain/10‡	Pre	4.58	4.74	4.78	71.80 (1,121) [<0.001†]	2.73 (2,121) [0.069]	6.11 (2,121) [<0.003†]
	Post	(2.50)	(2.48)	(2.56)			
		1.61	3.29	3.56			
		(1.83)	(2.71)	(3.13)			
		N = 41	N = 42	N = 41			
RMDQ–total no. of yes/24	Pre	17.85	16.69	18.37	64.10 (1,197) [<0.001†]	2.88 (2,199) [0.059]	12.19 (2,197) [<0.001†]
	Post	(5.66)	(5.97)	(5.34)			
		11.37	15.44	15.45			
		(7.55)	(6.92)	(7.34)			
Medication–total pills/day§	Pre	3.37	3.29	2.65	24.09 (1,71) [<0.001†]	1.04 (2,72) [0.357]	4.35 (2,71) [0.016†]
	Post	(2.92)	(2.74)	(2.38)			
		0.81	2.57	1.73			
		(2.25)	(2.77)	1.73)			
		N = 27	N = 21	N = 26			
During the past week, pain interference with usual work, in and outside home Rating (0–5)	Pre	3.41	3.49	3.39	89.80 (1,197) [<0.001†]	4.38 (2,197) [0.014†]	10.32 (2,197) [<0.001†]
	Post	(1.10)	(1.05)	(0.92)			
		2.24	3.06	2.88			
		(0.92)	(1.09)	(1.15)			
BDI-11	Pre	9.10	9.14	7.69	57.30 (1,195) [<0.001†]	1.31 (2,195) [0.273]	4.86 (2,195) [0.009†]
	Post	(8.13)	(8.74)	(6.98)			
		4.94	7.65	5.25			
		(6.11)	(8.50)	(5.25)			
Taking medication (yes)	Pre	48	35	39	McNemar test used for time effect		
	Post	(63.2%)	(51.5%)	(54.9%)	Matched $P < 0.001$†		
		7 (14.9%)	15 (50.0%)	14 (41.2%)	Opposite $P = 0.250$		
					EBC $P = 0.125$		

*Evidence-based care group.
†Significant at 0.05 level.
‡Those with no leg pain on admission are excluded from this analysis.
§Those with zero medication excluded from analysis.

of the 12 (41%) contacted by telephone stated that their dropout (Figure 1) was related to unrelieved or worsened pain. Only 3 of 16 baseline variables were statistically different between dropouts and those supplying complete data (Table 5). Dropouts were more likely to be female (65.5% *vs.* 44.9%), have a higher mean depression score (14.86 *vs.* 8.57), and have a lower percentage with prior episodes (53.6% *vs.* 74%). After dropouts, 201 subjects were eligible for analysis.

Complications Resulting in Early Withdrawal

As opposed to dropouts, "withdrawals" were unable/unwilling to continue their exercises for the full 2-week protocol (N = 36) because of no improvement, worsening, or increased distal radiation of symptoms. They provided their 2-week data early so they could move to alternative care. There was considerable variability in these withdrawal rates in that no matched group subjects withdrew compared with 16 (34.8%) of the opposite and 20 (32.8%) of the EBC groups ($\chi^2 = 18.67$, $df = 2$, $P < 0.001$). Withdrawal subjects were included in our analysis, which accounts for the lower number of visits for the opposite (2.81) and EBC (3.05) groups compared with the matched group (4.08) ($F = 16.15$, $df = 2203$, $P < 0.001$).

Treatment Effects

Sixty-eight percent of our 201 subjects returned their compliance questionnaires. Their overall compliance rating was good (3–4 sets/day) that did not differ significantly across treatment groups ($P = 0.121$).

All three treatment groups improved in all outcome measures over the 2-week trial. However, there was statistically significantly greater improvement in every outcome variable for the matched group compared with the opposite or EBC groups, with *P* values ranging from 0.016 to <0.001 (Table 6; and Figures 3, 4).

Satisfaction With Care

The matched group had statistically significantly greater improvement than the other two treatment groups (*P* values <0.005) in all five areas of satisfaction with care: return to work, home and recreational activity, perceived need for further treatment, and self-rated improvement (Table 7; Figure 3). The opposite group consistently faired the poorest on all satisfaction parameters; indeed, 15% to 17% of both the opposite and the EBC groups reported worsening, despite their expected favorable prognosis (Figure 3).

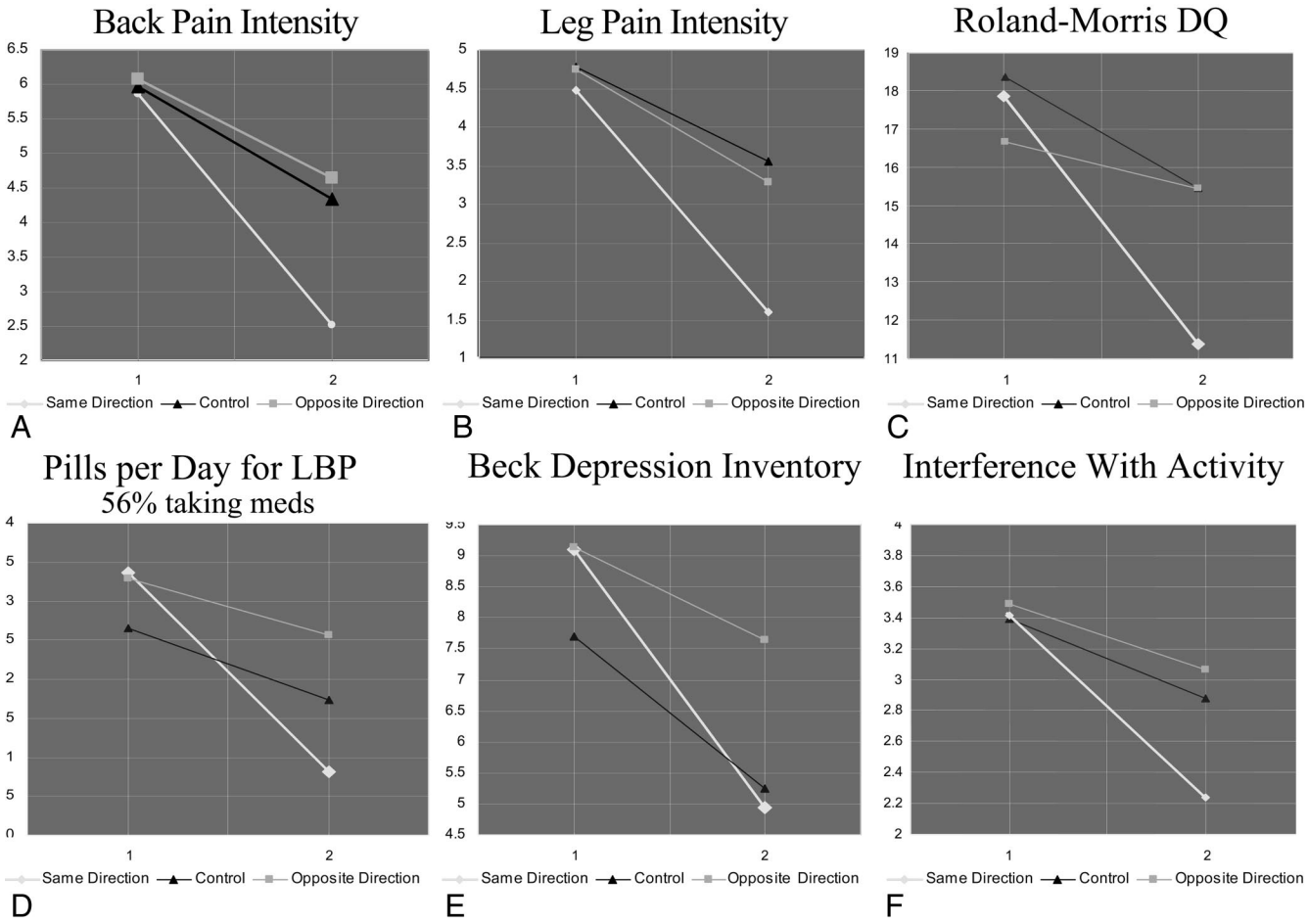

Figure 3. Outcomes. **A,** back pain (1–10) $P < .001$. **B,** leg pain (0–10) $P < .003$. **C,** Roland-Morris Disability Questionnaire (0–24) $P < 0.01$. **D,** medication total pills/day for those taking medication at baseline (56%) $P = .016$. **E,** Beck Depression Inventory (0–21) $P = .009$. **F,** activity interference (0–5) $P < .001$.

Changes in QTF Severity Classification

More than one third (36.8%) of the matched group reported improvement in QTF severity (pain location/neurologic status) classification compared with only 10.6% of the opposite and 19.3% of the EBC groups. Further, no matched group subject reported any deterioration in their severity classification compared with 12.8% of the opposite, and 17.5% of the EBC groups ($\chi^2 = 20.70$, $df = 4$, $P < 0.001$).

■ Discussion

Many trials describe exercise protocols extending over many visits, sometimes for months, suggesting the intent to address deconditioning.[6,46–48] Alternatively, patient-specific DP exercises have a specific focus: pain control and/or elimination.[23] Not surprisingly, gaining control over pain that is inhibiting function improved every other outcome measure in only four visits. Whether progression to cardiovascular or strength straining programs can further enhance these outcomes is worthy of further investigation.

Consistent with current evidence-based guidelines, the MDT method used in this study routinely provides activity advice and patient education, but these are based on DP principles. By study design, the mobilization and manipulation components of the MDT method were not included in the treatment protocol.[23]

Another documented contribution of this form of assessment is the ability to predict chronic pain and disability at 1 year[54] by the early identification of those who fail to demonstrate DP/centralization ("noncentralizers").

Based on the many studies reporting high reliability and predictive validity, this MDT form of assessment was given the highest grade for scientific support as

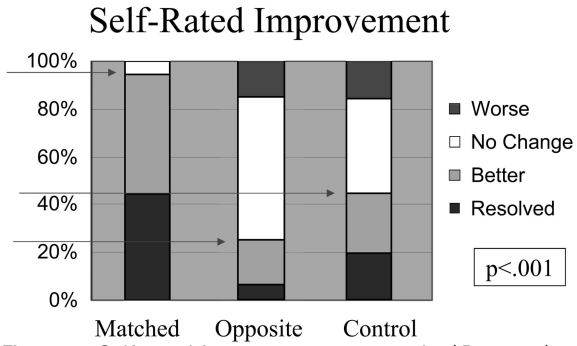

Figure 4. Self-rated improvement at 2 weeks ($P < .005$).

Table 7. Satisfaction With Care/Beliefs Regarding Need for Further Treatment

Satisfaction/Beliefs	Matched	Opposite	EBC*	χ^2 (*df*) [*P*]
Please tick the phrase that best describes how you feel you have responded to treatment:				
My back problem has resolved	32 (44.45)	4 (6.3%)	13 (19.1%)	74.63 (6) [<0.001†]
My back problem has improved	36 (50.0%)	12 (19.0%)	17 (25.0%)	
My back has not changed	4 (5.6%)	38 (60.3%)	27 (39.7%)	
My back problem is worse	0 (0.0%)	9 (15.3%)	11 (16.2%)	
Please record your wishes for further treatment:				
I do not need further treatment	30 (42.3%)	2 (3.5%)	10 (12.7%)	85.87 (4) [<0.001†]
I need a few more session of the same treatment	31 (43.7%)	3 (5.3%)	14 (17.7%)	
I prefer to change treatment to try to achieve better pain control	10 (14.1%)	52 (91.2%)	55 (69.7%)	
Regarding work and/or household chores, I feel I am able to:				
Return to all my previous activities.	28 (38.9%)	12 (19.0%)	14 (20.9%)	18.69 (6) [0.005†]
Return to all of my activities with minor adjustments.	21 (29.2%)	18 (28.6%)	13 (19.4%)	
Return to most but not all of my activities	18 (25.0%)	16 (25.4%)	26 (38.8%)	
Return to few or none of my activities	5 (6.9%)	17 (27.0%)	14 (20.9%)	
Regarding sports and recreational activities:				
I can participate in the same activities as before	28 (40.0%)	10 (15.9%)	11 (16.7%)	28.09 (6) [<0.001†]
I can participate in same activities with *minor* adjustments	26 (37.1%)	20 (31.7%)	27 (40.9%)	
I will have to change activities or make *major* adjustments	13 (18.6%)	23 (36.5%)	11 (16.7%)	
I will be unable to participate in sports and recreation.	3 (4.2%)	10 (15.9%)	16 (25.8%)	
Regarding work status, I will be returning to:				
Same job, full time, full duties	53 (76.8%)	26 (58.1%)	30 (44.8%)	26.03 (10) [0.004†]
Same work, full time, modified	4 (5.8%)	5 (8.1%)	10 (14.9%)	
Same work, part time	5 (7.2%)	9 (14.5%)	4 (6.0%)	
I will have to find different work because of back pain	1 (1.4%)	0 (0.0%)	4 (6.0%)	
I am unable to work because of back pain	3 (4.3%)	0 (14.5%)	10 (14.9%)	
Other	3 (4.3%)	3 (4.8%)	9 (13.4%)	

*Evidence-based group.
†Significant at <0.05 level.

"both a diagnostic tool and a prognostic indicator" by Denmark's LBP Clinical Guidelines.[50] Our results add even further validation to that recommendation.

This is the first study that uses subgroups based solely on DP/centralization and directly links this subgroup's outcome to a subgroup-specific treatment (*i.e.*, individual exercise prescription based on DP). Previous studies reporting a good prognosis in this subgroup had not tested the role of treatment selection in the better outcomes.[32–35] In contrast, other commonly used classifications, such as QTF 1, 2, 3, and 4, have shown little clinical value beyond being descriptors of the population.[51,52,58] Further, commonly used labels such as "acute" and "chronic" only describe the current episode while failing to recognize the recurrent nature of the condition. Our high rate of prior episodes (70%), consistent with others' findings, illustrates the limitations of using the simplistic descriptors of acute, subacute, and chronic.[49,53] Further, 46.2% of our DP subjects were considered chronic and by conventional wisdom would be considered more difficult to treat; yet, as in many other studies of centralization and DP,[29,31–35] good and excellent outcomes are the rule when a pain-controlling exercise prescription matches individual's DP.

One of our treatment groups performed exercises in the direction opposite to their baseline DP. While this group's mean pain ratings and Roland Morris Disability Questionnaire scores improved, 34.8% of this group withdrew early, 60.3% reported no improvement, and 15.3% reported worsening (Table 7, Figure 3). Six subjects (12.8%) also reported "worsening" in their QTF severity rating, *i.e.*, their pain was farther down their leg at 2 weeks than at baseline.

Given the many studies reporting good to excellent prognoses for patients demonstrating a DP and centralization,[15,32–35,54–56] most of our study participants would be expected to improve regardless of treatment, based on being active, exercising, natural history, placebo and Hawthorne effects, and regression to the mean. But even the EBC group, who exercised daily and received both advice to remain active and reassurance of likely recovery, did not fare nearly as well as anticipated and significantly less well than those treated with the matched exercises. Fifty-six percent of the EBC group reported they were overall no better or were worse (Table 7), with 17.5% reporting pain radiating further into their leg after 2 weeks of treatment. The outcomes of many EBC subjects, in whom improvement was anticipated, were ap-

parently jeopardized by a counterproductive exercise prescription.

That not a single subject in the matched group demonstrated overall worsening or pain further into their leg (Table 7) is not only an important finding; it also reflects the overall safety of the DP-matched exercises when applied to patients in whom a DP was elicited by a properly trained clinician.

Of our original 312 subjects who underwent the MDT assessment, 53.5% demonstrated a DP for pure sagittal extension, but the remainder either required a hips-off-center start position or movements in other planes to centralize/abolish their pain, or had no DP elicited. This refutes the commonly held myth that the MDT (McKenzie) method consists only of extension exercises.[4,57]

Strengths and Limitations

The authors hope our descriptions of our study sample and treatment procedures will encourage replication of this study.[40] No special recruitment techniques were used to draw subjects who would not otherwise seek care. Withdrawals were analyzed in an intention-to-treat analysis, and dropouts are well documented.

Our primary limitation is the potential bias introduced by the opposite and EBC treatments being provided by MDT-trained PTs. However, we think such bias is minimal, if any, since all treating PTs were trained in all three treatment interventions and were instructed to confidently and enthusiastically present all exercises based on the expectation that all three groups would improve, as previously discussed. All PTs were also motivated to prevent dropouts and maintain the power of the study.

As in other randomized controlled trials, care-seeking subjects precluded the feasibility of using a "no treatment" control group, and neither subjects nor treating PTs could be blinded to their treatment.[4,47] Subjects were, however, shielded from their directional preference identified at baseline. The large number of subjects unable or unwilling to complete the full 2 weeks of treatment confirmed the pragmatic and ethical impossibility of maintaining the randomized treatments for a longer period.

A phenomenon of potential importance to this and many other LBP clinical studies is illustrated by reviewing the first subject randomized to the opposite treatment group (classified as an extension DP assigned to flexion exercises) who reported a 2-point improvement in her pain intensity and then returned to modified work duties. A post-study interview revealed that, when not performing her flexion exercises, she avoided forward bending and sitting (flexion) because they consistently increased her pain, leading her to do more standing and walking (extension activities) than usual, simply because this "felt better." Given her assignment to flexion treatment, this self-imposed activity modification of avoiding flexion activities and positions in favor of extension is a form of directional "contamination" of her treatment and may well have been more influential in her final outcome, *i.e.*, overall improvement, despite her actual study assignment of potentially aggravating flexion exercises. Future research might do well to focus on the role of subject's directional self-modification of daily activities in response to pain monitoring that would influence the outcome of any therapeutic intervention or recovery by "natural history."

■ Conclusion

Exercises concordant with patients' DP significantly improved outcomes compared with nonconcordant exercises and advice, and appear to be an effective pain control/elimination treatment strategy. This refutes prior systematic reviews concluding that specific exercises are not warranted.

While highly favorable outcomes are well documented in many studies for the DP subgroup of LBP, this study adds further validity by demonstrating that a subject-specific treatment is superior to others in creating good outcomes for this subgroup.

Further, the literature-predicted favorable prognosis for the DP subgroup was actually compromised by our version of "evidence-based care." Our data also introduce evidence for the existence of counterproductive exercises in many LBP patients.

■ Key Points

- A subgroup defined by the presence of a directional preference has better outcomes when prescribed exercises that match the individual's directional preference than with nonindividualized evidence-based care.
- The good prognosis documented in other studies for study participants with directional preference and centralization findings at baseline assessment is significantly diminished if exercise prescriptions do not match their directional preference findings.
- Reliable identification of a previously validated low back pain subgroup before our randomization provided homogeneity to our study sample, which likely contributed to the significant differences found between treatment group outcomes.

References

1. Faas A. Exercises: which ones are worth trying, for which patients, and when? *Spine* 1996;21:2874–2878; discussion 2878–2879.
2. Koes B, Boutar L, Beckerman H, et al. Physiotherapy exercises and back pain: a blinded review. *Br Med J* 1991;302:1572–1576.
3. Spitzer W, LeBlanc F, Dupuis M. Scientific approach to the assessment and management of activity-related spinal disorders (the Quebec Task Force). *Spine* 1987;12(suppl):16–21.
4. van Tulder M, Malmivaara A, Esmail R, et al. Exercise therapy for low back pain. *Spine* 2000;25:2784–2796.
5. van Tulder M, Koes B, Bouter L. Conservative treatment of acute and chronic nonspecific low back pain: a systematic review of randomized controlled trials of the most common interventions. *Spine* 1997;22:2128–2156.
6. Abenhaim L, Rossignol M, Valat JP, et al. The role of activity in the therapeutic managment of back pain: report of the International Paris Task Force on Back Pain. *Spine* 2000;25(suppl):1–33.

7. Bigos S, Bowyer O, Braen G, et al. *Acute Low Back Problems in Adults: Clinical Practice Guidelines* [quick reference guide no. 14]. Rockville, MD: U.S. Department of Health and Human Services, Agency for Health Care Policy and Research, 1994.
8. Practitioners RCoG. Clinical guidelines for the management of acute low back pain. *London: Royal College of General Practitioners*, 1996, 1999.
9. Battie M, Cherkin D, Dunn R, et al. Managing low back pain: attitudes and treatment preferences of physical therapists. *Phys Ther* 1994;74:219–226.
10. Foster N, Thompson K, Baxter D, et al. Management of nonspecific low back pain by physiotherapists in Britain and Ireland. *Spine* 1999;24:1332–1342.
11. Hsieh C, Adams A, Tobis J, et al. Effectiveness of four conservative treatments for subacute low back pain: a randomized clinical trial. *Spine* 2002; 27:1142–1148.
12. Dettori JR, Bullock SH, Sutlive TG, et al. The effects of spinal flexion and extension exercises and their associated postures in patients with acute low back pain [see comments]. *Spine* 1995;20:2303–2312.
13. Elnaggar I, Nordin M, Sheikhzadeh A, et al. Effects of spinal flexion and extension exercises on low-back pain and spinal mobility in chronic mechanical low-back pain patients. *Spine* 1991;16:967–972.
14. Koes B, Bouter L, van der Heijden G. Methodological quality of randomized clinical trials on treatment efficacy in low back pain. *Spine* 1995;20:228–235.
15. Delitto A, Cibulka M, Erhard R, et al. Evidence for an extension-mobilization category in acute low back syndrome: a prescriptive validation pilot study. *Phys Ther* 1993;73:216–228.
16. Erhard R, Delitto A, Cibulka M. 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: 1093–1100.
17. Fritz J, George S. The use of a classification approach to identify subgroups of patients with acute low back pain. *Spine* 2000;25:106–114.
18. Kopp JR, Alexander AH, Turocy RH, et al. The use of lumbar extension in the evaluation and treatment of patients with acute herniated nucleus pulposus: a preliminary report. *Clin Orthop* 1986;202:211–218.
19. O'Sullivan P, Twomey L, Allison G. Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylosis or spondylolisthesis. *Spine* 1997;22:2959–2967.
20. Spratt K, Lehrmann T, Weinstein J, et al. A new approach to the low back physical examination: behavioral assessment of mechanical signs. *Spine* 1990;15:96–102.
21. Borkan J, Koes B, Reis S, et al. A report from the second international forum for primary care research on low back pain: reexamining priorities. *Spine* 1998;23:1992–1996.
22. Bouter L, Pennick V, Bombardier C. Cochrane back review group. *Spine* 2003;28:1215–1218.
23. McKenzie R, May S. *Mechanical Diagnosis and Therapy*, 2nd ed. Waikanae, New Zealand: Spinal Publications New Zealand Ltd., 2003.
24. Fritz J, Delitto A, Vignovic M, et al. Interrater reliability of judgments of the centralization phenomenon and status change during movement testing in patients with low back pain. *Arch Phys Med Rehabil* 2000;81:57–61.
25. Kilby J, Stigant M, Roberts A. The reliability of back pain assessment by physiotherapists, using a 'McKenzie algorithm.' *Physiotherapy* 1990;76: 579–583.
26. Kilpikoski S, Airaksinen O, Kankaanpää M, et al. Interexaminer reliability in low back pain assessment using the McKenzie method. *Spine* 2002;27:E207–E214.
27. Razmjou H, Kramer J, Yamada R. Inter-tester reliability of the McKenzie evaluation of mechanical low back pain. *J Orthop Sports Phys Ther* 2000; 30:368–383.
28. Spratt K, Weinstein J, Lehmann T, et al. Efficacy of flexion and extension treatments incorporating braces for low-back pain patients with retrodisplacement, spondylolisthesis, or normal sagittal translation. *Spine* 1993;18: 1839–1849.
29. Werneke M, Hart DL, Cook D. A descriptive study of the centralization phenomenon: a prospective analysis. *Spine* 1999;24:676–683.
30. Wilson L, Hall H, McIntosh G, et al. Intertester reliability of a low back pain classification system. *Spine* 1999;24:248–254.
31. Donelson R, Aprill C, Medcalf R, et al. A prospective study of centralization of lumbar and referred pain: a predictor of symptomatic discs and anular competence. *Spine* 1997;22:1115–1122.
32. Donelson R, Silva G, Murphy K. The centralization phenomenon: its usefulness in evaluating and treating referred pain. *Spine* 1990;15:211–213.
33. Karas R, McIntosh G, Hall H, et al. The relationship between nonorganic signs and centralization of symptoms in the prediction of return to work for patients with low back pain. *Phys Ther* 1997;77:354–360.
34. Long A. The centralization phenomenon: its usefulness as a predictor of outcome in conservative treatment of chronic low back pain. *Spine* 1995;20: 2513–2521.
35. Sufka A, Hauger B, Trenary M, et al. Centralization of low back pain and perceived functional outcome. *J Orthop Sports Phys Ther* 1998;27:205–212.
36. Donelson R, Grant W, Kamps C, et al. Pain response to end-range spinal motion in the frontal plane: a multi-centered, prospective trial. *International Society for the Study of the Lumbar Spine*, Heidelberg, Germany, 1991.
37. Donelson R, Grant W, Kamps C, et al. Pain response to repeated end-range sagittal spinal motion: a prospective, randomized, multi-centered trial. *Spine* 1991;16:206–212.
38. Snook S, Webster B, McGorry R, et al. The reduction of chronic nonspecific low back pain through the control of early morning lumbar flexion: a randomized controlled trial. *Spine* 1998;23:2601–2607.
39. Williams M, Hawley J, McKenzie R, et al. A comparison of the effects of two sitting postures on back and referred pain. *Spine* 1991;16:1185–1191.
40. Deyo RA, Battie M, Beurskens AJ, et al. Outcome measures for low back pain research: a proposal for standardized use. *Spine* 1998;23:2003–2013.
41. Indahl A, Velund L, Reikeraas O. Good prognosis for low back pain when left untampered: a randomized clinical trial. *Spine* 1995;20:473–477.
42. Riddle D, Rothstein J. Intertester reliability of McKenzie's classifications of the syndrome types present in patients with low-back pain. *Spine* 1994;18: 1333–1344.
43. Jensen M, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. *Pain* 1986;27:117–126.
44. Roland M, Morris R. A study of the natural history of low back pain: Part 1. Development of a reliable and sensitive measure of disability in low back pain. *Spine* 1983;8:141–144.
45. Beck A, Steer R, Brown, G. *Beck Depression Inventory*, vol. II, Manual. San Antonio, TX: Psychological Corporation, 1996.
46. Beurskens AJ, de Vet HC, Koke AJ, et al. Measuring the functional status of patients with low back pain. *Spine* 1995;20:1017–1028.
47. Mannion A, Müntener M, Taimela S, et al. A randomized clinical trial of three active therapies for chronic low back pain. *Spine* 1999;24:2435–2448.
48. Soukup MG, Glomsrod B, Lonn JH, et al. The effect of a Mensendieck exercise program as secondary prophylaxis for recurrent low back pain. *Spine* 1999;24:1585–1591.
49. Von Korff M, Ormel J, Keefe F, et al. Grading the severity of chronic pain. *Pain* 1992;21:2833–2839.
50. Danish Institute for Health Technology Assessment. *Low Back Pain: Frequency, Management, and Prevention From an HTA Perspective*. Danish Health Technology Assessment 1999;1(1).
51. Loisel P, Vachon B, Lamaire J, et al. Discriminiative and predictive validity assessment of the Quebec Task Force classification. *Spine* 2002;27:851–857.
52. Werneke M, Hart D. Categorizing patients with occupational low back pain by use of the Quebec Task Force classification system versus pain pattern classification procedures: discriminant and predictive validity. *Phys Ther* 2004;84:243–254.
53. McGorry R, Webster B, Snook S, et al. The relation between pain intensity, disability, and the episodic nature of chronic and recurrent low back pain. *Spine* 2000;25:834–841.
54. Werneke M, Hart DL. Centralization phenomenon as a prognostic factor for chronic low back pain and disability. *Spine* 2001;26:758–765.
55. Werneke M, Hart D. Discriminant validity and relative precision for classifying patients with nonspecific neck and back pain by anatomic pain patterns. *Spine* 2003;28:161–166.
56. Aina S, May S, Clare H. The centralization phenomenon of spinal symptoms: a systematic review. *Man Ther* 2004;9:134–143.
57. Donelson R. Letter to the editor. Spine 2001;26:1827–1829.
58. Padfield B, Chesworth B, Butler R. Use of an outcome measurement system to answer a clinical question: Is the Quebec Task Force classification system useful in an outpatient setting? Physiotherapy Canada 2002;54(4):258–64.