SPINE Volume 32, Number 16, pp 1785–1790
©2007, Lippincott Williams & Wilkins, Inc.

Statistical Significance *Versus* Clinical Importance

Trials on Exercise Therapy for Chronic Low Back Pain as Example

Maurits van Tulder, PhD,*†§ Antti Malmivaara, MD, PhD,‡ Jill Hayden, DC,§
and Bart Koes, PhD||

Study Design. Critical appraisal of the literature.

Objecives. The objective of this study was to assess if results of back pain trials are statistically significant and clinically important.

Summary of Background Data. There seems to be a discrepancy between conclusions reported by authors and actual results of randomized controlled trials. Little attention has been paid to the problem of over-reporting of conclusions.

Methods. All 43 trials of the Cochrane review on exercise therapy for low back pain were included. Descriptive analyses were conducted.

Results. Eighteen trials reported positive conclusions in favor of exercise. Only six of the 43 studies showed both clinically important and statistically significant differences in favor of the exercise groups on function, and 4 on pain.

Conclusion. It seems that many conclusions of studies of exercise therapy for chronic low back pain have been based on statistical significance of results rather than on clinical importance and, consequently, may have been too positive. Authors of trials should report not only statistical significance of results but also clinical importance.

Key words: low back pain, randomized controlled trials, statistical significance, clinical importance. **Spine 2007;32: 1785–1790**

In conducting systematic reviews in the field of low back pain, the authors have noticed that there is a discrepancy between conclusions reported by authors and actual results of randomized controlled trials (RCTs). Authors' conclusions are usually too positive. Aspects of the design and conduct of RCTs as well as aspects of the reporting of results may affect the conclusions concerning effectiveness of the intervention at issue. Little attention has been paid to this problem of over-reporting of conclusions.

From the *Institute for Research in Extramural Medicine, VU University Medical Center, Amsterdam, the Netherlands; †Department of Health Economics & Health Technology Assessment, Institute of Health Sciences, Faculty of Earth and Life Sciences, VU University, Amsterdam, the Netherlands; ‡Finnish Office for Health Care Technology Assessment, FinOHTA/Stakes, Helsinki, Finland; §Institute for Work & Health, Toronto, Ontario, Canada; and ||Department of General Practice, Erasmus Medical Center, Rotterdam, the Netherlands.

Acknowledgment date: June 16, 2006. First revision date: December 12, 2006. Acceptance date: January 12, 2007.

The manuscript submitted does not contain information about medical device(s)/drug(s).

Institutional funds were received in support of 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 Maurits van Tulder, PhD, Institute for Research in Extramural Medicine, VU University Medical Center, Van der Boechorststraat 7, 1081 BT Amsterdam, the Netherlands; E-mail: mw.van_tulder.emgo@med.vu.nl

Many papers have been published on validity of trial results, which is defined by the methodologic quality of the trial. Suggestions have been made to improve the reporting of trials enabling assessment of methodologic quality, for example, the CONSORT statement.[1] RCTs with high methodologic quality are less likely to provide biased results than trials with poor methodologic quality.

Authors of RCTs usually report statistically significant differences between groups, and conclusions are often based on this statistical significance.[2] If $P < 0.05$, the conclusion is usually that the intervention is more effective than the comparison; if $P > 0.05$, the intervention is considered not more effective. This P value is not very informative and only indicates the chance of the observed effect, not considering its size.[3] The P value does not indicate if the effect is clinically important. Clinical importance is defined by the minimally clinically important difference (MCID) between groups.[4]

Multiple outcomes are used in most back pain trials. A core set of outcome measures has been recommended for back pain trials, including pain, function, generic health status, and work disability.[5] Few trials, however, have evaluated the MCID on these outcome measures. Only some trials evaluated the MCID for function[6–11] and for pain[11,12] in low back pain populations.

The objective of this paper was to assess if treatment effects of low back pain trials are statistically significant and clinically important.

Methods

All 43 RCTs on chronic low back pain (defined as low back pain for 12 weeks or more) that were included in the recent update of the Cochrane review on exercise therapy for low back pain were included in this study.[13] Two reviewers (J.H., M.vT.) independently assessed the quality and extracted relevant data from these studies. Any discrepancies were discussed in a consensus meeting. If disagreement persisted, a third reviewer was consulted, who made a final decision taking into account the arguments of both reviewers.

To assess statistical significance and clinical importance, data were extracted on the magnitude of effect on primary outcome measures for all reported follow-up moments. We classified follow-up data into earliest follow-up, short-term (closest to 6 weeks), intermediate (closest to 6 months), and long-term (closest to 12 months) follow-up. These data were used in the meta-analysis conducted in the Cochrane review.[13] Within-group changes were either extracted from the reports of the studies or calculated. Between-group differences on primary outcome measures were based on the difference in change scores compared to baseline, and also extracted from the reports of the studies or calculated. If multiple time periods or multiple exercise and comparison groups were available, we used the largest difference from the trial.

Statistical Significance. Between-group differences were considered statistically significant at the 5% level.

Clinical Importance. Between-group differences were compared with observations on the MCID reported in the literature.[6,9,11,12] Means and standard deviations for baseline and all follow-up moments were extracted and transformed to a 0 to 100 scale. Between-group differences were considered clinically important if the magnitude was 20% or more for pain and 10% or more for functioning. Because only a few studies included data on work absenteeism and global improvement, we did not evaluate the within-group changes and between-group differences for these 2 outcomes.

Results

Thirty-seven of the 43 RCTs reported on pain (Table 1).[15–52] Sixteen of the 37 RCTs (43%) showed statistically significant differences on pain in favor of the exercise group,[16,17,23,24,26,30,33,37–40,42,46–48,50] and 2 were unclear.[22,35] Only six of the 37 RCTs (16%) showed clinically important between-group differences on pain,[17,22,33,35,40,50] and 31 did not.

Thirty-four of the 43 RCTs reported on function (Table 1).[16–19,21–24,26–28,30–46,48–54] Fifteen of the 34 RCTs (44%) showed statistically significant differences between groups on function in favor of the exercise group,[16,17,23,26,33,36–40,42,46,48,50,53] and 2 were unclear.[22,35] Seven of the 34 trials (21%) showed clinically important between-group differences on function,[16,17,22,33,37,40,42] and 2 were unclear.[24,53]

Four of the 37 RCTs (11%) reporting on pain showed both clinically important and statistically significant differences on pain in favor of the exercise group compared with the nonexercise control group.[17,33,40,50]

Six of the 34 RCTs (18%) reporting on function showed both clinically important and statistically significant differences in favor of the exercise groups compared with another type of exercise[16,17] or with a nonexercise control group.[33,37,40,42] Four RCTs reported on neither pain nor function.[55–58]

Discussion

The results of this study show that 18 trials on chronic low back pain reported positive conclusions in favor of exercise, but only 7 showed clinically important differences.

Findings were similar for the studies on acute and subacute low back pain (data not presented). Positive conclusions should be based on statistical significant and clinically important outcomes. Small studies may show clinically important findings that are not statistically significant. A meta-analysis may resolve this problem of underpowered studies. However, if findings are statistically significant but not clinically important, a meta-analysis will not change the conclusions. The meta-analysis of the Cochrane review on exercise therapy for low back pain found that exercise therapy appears to be effective at decreasing pain and improving function slightly in adults with chronic low back pain, particularly in healthcare populations.

Our study has some limitations. First, relatively little empirical evidence exists on MCID of low back pain outcomes. We used data reported by Hägg *et al*, but these came from a population of severe chronic patients with low back pain referred to surgery.[12] This population may have been different from the patient populations included in exercise trials. The thresholds that we used of 20% for pain and 10% for function were arbitrary and not very strict. Three other studies found a MCID for function in patients with low back pain of 10% to 20%[6,8,9] and one found a MCID for pain of more than 30%.[11] If we would have used higher thresholds for the MCID, for example, 30% for pain and 20% for function, none of the trials would have shown clinically important differences on pain or function between groups. Future studies are direly needed that evaluate the MCID for the most important outcome measures in low back pain research.

The use of multiple outcomes in the field of low back pain research warrants caution. If the intervention of interest is statistically significantly better than the comparison on one of the outcome measures, the conclusion usually is that the intervention is more effective than the comparison regardless if there is no difference on any of the other outcome measures. Obviously, one would expect that the results support a prior hypothesis about the potential working mechanisms of the intervention. These mechanisms should *a priori* lead to the choice of one primary outcome measure. This primary outcome measure should be used to define the expected difference in effect in sample size calculations. The expected difference should be at least as large as the MCID.

Pocock *et al* already in 1987 emphasized that reports of RCTs are biased toward an exaggeration of treatment effects and that more emphasis should be given to the magnitude of the effect.[59] This study shows that studies on exercise therapy for low back pain have focused on statistical significance and have overestimated treatment effects. Eighteen trials on chronic low back pain reported positive conclusions in favor of the exercise group, while only 7 of these studies showed clinically important differences. The fact that only a few exercise therapy studies showed clinically important effects is disappointing for an intervention that is considered to be effective and is recommended in clinical guidelines for subacute and chronic low back pain. We argue that other interventions that seem to be effective for low back pain, such as advice to stay active and nonsteroidal anti-inflammatory drugs for acute low back pain, and behavioral therapy and multidisciplinary treatment for chronic low back pain probably show similar sobering results. Future studies should evaluate the clinical importance of effects for other low back pain interventions.

Little attention has been paid to adequacy of reporting of results and conclusions of back pain studies. Although we were able to assess the clinical importance of most

Table 1. Outcome Measures Reported and Within-Group Change and Between-Group Difference on Pain and Functioning in Trials on Exercise Therapy for Chronic Low Back Pain

		Clinical Importance of Within-Group Changes From Baseline				Clinical Importance and Statistical Significance of Between-Group Changes			
		Pain		Function		Pain		Function	
Study	Exercise Group	Exercise	Control	Exercise	Control	Importance	Significance	Importance	Significance
Alexandre (2001)[15]	Multiple component exercises	Y	Y			N	N		
Aure (2003)[16]	Stretch two thirds, passive manipulation (one third)	Y		Y		N	Y	Y	Y
	Individually designed exercise program	N		Y					
Bendix (1995)[17]	Functional restoration	Y		Y		Y	Y	Y	
	Aerobics, strengthening	N		N					
Bendix (2000)[18]	Aerobics, strengthening	N	N	Y	Y	N	N	N	N
	Functional restoration								
Bentsen (1997)[53]	Dynamic strengthening			?	?			?	Y
	Home exercises								
Bronfort (1996)[19]	Dynamic trunk (Manniche)	Y	Y	Y	Y	N	N	N	N
	Same plus NSAIDs	Y		Y		N	N	N	N
Buswell (1982)[20]	Extension	Y				N	N		
	Flexion	Y							
Calmels (2004)[21]	Isokinetic strengthening (Cybex)	N		N		N	N	N	N
	Physiotherapy exercises	N		N		N	N	N	N
Dalichau (2000)[22]	Strengthening with lumbar support	Y	N	Y	N	Y	?	Y	?
	Same with no lumbar support								
Descarreaux (2002)[23]	Standard strengthening, stretching	N		Y		N	Y	N	Y
	Force, extensibility exercises	N		N					
Deyo (1990)[24]	Relaxation, stretching exercises	Y	N	?	?	N	Y	?	N
Elnaggar (1991)[25]	Extension	N				N	N		
	Flexion	N							
Frost (1995)[26]	Stretching, progressive exercises	N	N	N	N	N	Y	N	Y
Frost (2004)[27]	Standard PT			N	N			N	N
Galantino (2004)[54]	Hatha yoga			?	?			?	?
Gur (2003)[28]	Stretching, strengthening	Y	Y	Y	Y	N	N	N	N
	Same with laser intervention	Y		Y					
Hansen (1993)[29]	Intensive dynamic training	N	N			N	N		
Hemmila (1997)[30]	Bending, rotation exercises	N	Y	N	N	N	Y*	N	Y*
			N		N		N	N	N
Hildebrandt (2000)[56]	Postural (Cesar therapy)								
Johanssen (1995)[31]	Aerobics, exercises emphasizing coordination	Y		Y		N	N	N	N
	Aerobics, exercises emphasizing endurance	Y		Y					
Jousset (2004)[32]	Functional restoration	N		Y		N	N	N	N
	Active individual therapy	N		N					
Kankaanpaa (1999)[33]	DBC program	Y	N	Y	N	Y	Y	Y	Y
Kendall (1968)[57]	Mobilizing, strengthening, posture								
	Isometric flexion								
	Strengthening extension								
Kuukkanen (2000)[34]	Strengthening, endurance, balance	N	N	Y	N	N	N	N	N
		N		N					
Lidstrom (1970)[58]	Mobilizing, strengthening								
Lie (1999)[59]	Mobilizing, stretching, walking in "flexible way"								
	Stabilizing exercises								
Manniche (1988)[35]	Intensive strengthening	Y		Y		Y	?	N	?
	Back strengthening (less dose)	N		Y		N	?	N	?
	Isometric exercises	N		Y					
Mannion (1999)[36]	Aerobics, strengthening	N		Y		N	N	N	Y
	Controlled progressive exercises with machines	N		Y		N	N	N	Y
	PT including exercises with Therabands	N		N					
Moseley (2002)[37]	Specific trunk muscle training	Y	N	Y	Y	N	Y	Y	Y
Niemisto (2003)[38]	Stabilizing exercises	Y	?	Y	Y	N	Y	N	Y
Petersen (2002)[39]	Strengthening training	Y		Y		N	Y	N	N
	McKenzie	N		Y		N	N	N	Y
Preyde (2000)[40]	Stretching, flexion-extension	N	N	N	N	Y	Y	Y	Y
	Comprehensive massage + exercise	Y	Y	Y	Y	Y	Y	Y	Y

(*Continued*)

Table 1. Continued

		Clinical Importance of Within-Group Changes From Baseline				Clinical Importance and Statistical Significance of Between-Group Changes			
		Pain		Function		Pain		Function	
Study	Exercise Group	Exercise	Control	Exercise	Control	Importance	Significance	Importance	Significance
Rasmussen-Barr (2003)[41]	Stabilizing exercises	N	N	Y	N	N	N	N	N
Risch (1993)[42]	Dynamic extension strengthening	N	N	N	N	N	Y	Y	Y
Rittweger (2002)[43]	Lumbar extension, resistance exercises	Y				N			
	Specific "oscillating platform" exercises	Y				N			
Soukup (1999)[44,45]	Mensendiek	N	N	Y	Y	N	N	N	N
Torstensen (1998)[46]	Ordinary activity level walking program	N	N	N	N	N	N	N	N
	MET	N	N	N	N	N	Y	N	Y
Tritilanunt (2001)[47]	Aerobics	Y	N			N	Y		
	Flexion exercises	N					N		
Turner (1990)[48]	Aerobic fitness	N	N	N	N	N	Y	N	Y
	Behavioural/exercises	N	N	Y	N	N	Y	N	Y
Yelland (2004)[49]	Sagittal flexibility	?	?	Y	Y	N	N	N	
Yeung (2003)[50]	Stretching, mobilizing	N		N		N	N	N	N
	Same plus electro-acupuncture	Y		N		Y	Y	N	Y
Yozbatiran (2004)[51]	Stretching, progressive exercises on land	Y		Y		N	N	N	N
	Same in water	Y		Y					
Zylbergold (1981)[52]	Flexion exercises	N	N	N	Y	N	N	N	N

Statistical significance at alpha $\geq$0.05. Clinical significance at 20% for pain and 10% for function.
*Effect of between-group differences in favor of control group.
†Qualitative, consensus-derived reviewer rating of overall conclusions about the effectiveness of the intervention.
Y indicates yes; N, no; ?, unclear due to insufficient data (effect of between group differences in favor of exercise group).

studies, there is still room for improvement. For example, one study reported that "lumbar pain in the last 7 days showed a *P* value of 7%, which could be attributed to the significance value in the 93% confidence interval."[15] Also, most studies did not report 95% confidence intervals around the point estimate of effect, but means and standard deviations. A confidence interval gives an estimated range of values (interval) for the variable of interest (*e.g.*, mean difference). It also quantifies the probability (confidence level) that the interval will include the true value. This confidence level is usually 95%.[60] Confidence intervals would be more informative and would facilitate assessment of clinical importance.[2] If the MCID is less than the lower limit of the 95% confidence interval, results are likely to be statistically significant and clinically important. If the MCID is greater than the upper limit of the 95% confidence interval, results are likely to be clinically unimportant. If the MCID lies within the limits of the 95% confidence interval, it is unclear if the effect is clinically important or not.[2] The reporting of results of studies on exercise therapy for low back pain should be improved, enabling readers to assess the clinical importance of the results. At present, it seems that many conclusions of studies of exercise therapy for low back pain have been based on statistical significance of results rather than on clinical importance and, consequently, may have been too positive. Journal editors and reviewers should be more alert in evaluating both statistical significance and clinical importance when judging if the conclusions of authors are supported by the results of trials.

Finally in accordance with the recommendations of Chan *et al*[2] that may lead to a reduction of the number of studies that exaggerate the results, we would like to suggest that authors: 1) choose one primary outcome measure for a trial, based on a plausible working mechanism of the intervention; 2) perform a sample size calculation; 3) use the MCID for the primary outcome measure as expected difference in sample size calculation; 4) report within-group changes and between-group differences; 5) report mean between-group differences with 95% confidence intervals; 6) check if the MCID used in the sample size calculation is smaller than the lower limit of the 95% confidence interval; and 7) only conclude that an intervention is more effective than the comparison group if No. 6 is fulfilled.

■ Key Points

- A discrepancy exists between conclusions reported by authors and actual results of randomized trials.
- Eighteen trials on chronic low back pain reported positive conclusions in favor of exercise, but only 6 of the 43 studies showed both clinically important and statistically significant differences in favor of the exercise groups on function, and 4 on pain.

- Reporting confidence intervals would facilitate assessment of clinical importance.
- Conclusions of trials should not only be based on statistical significance of effects, but also on clinical importance.

References

1. Moher D, Schulz KF, Altman DG, et al. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001;357:1191–4.
2. Chan KBY, Man-Son-Hing M, Molnar FJ, et al. How well is the clinical importance of study results reported? An assessment of randomized controlled studies. *CMAJ* 2001;165:1197–202.
3. Goodman SN. Toward evidence-based medical statistics: Part 1. The P value fallacy. *Ann Intern Med* 1999;130:995–1004.
4. Beaton DE, Boers M, Wells GA. Many faces of the minimal clinically important difference (MCID): a literature review and directions for future research. *Curr Opin Rheumatol* 2002;14:109–14.
5. Deyo RA, Battie M, Beurskens AJ, et al. Outcome measures for low back pain research: a proposal for standardized use. *Spine* 1998;23:2003–13.
6. Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain* 1996;65:71–6.
7. Fritz JM, Irrgang JJ. A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale. *Phys Ther* 2001;81:776–88.
8. Riddle DL, Stratford PW, Binkley JM. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: Part 2. *Phys Ther* 1998;78:1197–207.
9. Stratford PW, Binkley JM, Riddle DL, et al. Sensitivity to change of the Roland–Morris Back Pain Questionnaire: Part 1. *Phys Ther* 1998;78:1186–96.
10. Davidson M, Keating JL. A comparison of five low back disability questionnaires: reliability and responsiveness. *Phys Ther* 2002;82:8–24.
11. Van der Roer N, Ostelo RW, Bekkering GE, et al. Minimal clinically important change for pain intensity, functional status, and general health status in patients with nonspecific low back pain. *Spine* 2006;31:578–82.
12. Hägg O, Fritzell P, Nordwall A. The clinical importance of changes in outcome scores after treatment for chronic low back pain. *Eur Spine J* 2003;12:12–20.
13. Hayden JA, van Tulder MW, Malmivaara A, et al. Meta-analysis: exercise therapy for non-specific low back pain. *Ann Intern Med* 2005;142:765–75.
14. Farrar JT, Young JP Jr, LaMoreaux L, et al. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* 2001;94:149–58.
15. Alexandre NM, de Moraes MA, Correa Filho HR, et al. Evaluation of a program to reduce back pain in nursing personnel. *Rev Saude Publica* 2001;35:356–61.
16. Aure O, Nilsen J, Vasseljen O. Manual therapy and exercise therapy in patients with chronic low back pain: a randomized, controlled trial with 1-year follow-up. *Spine* 2003;28:525–32.
17. Bendix AF, Bendix T, Ostenfeld S, et al. Active treatment programs for patients with chronic low back pain: a prospective randomized, observer-blinded study. *Eur Spine J* 1995;4:148–52.
18. Bendix T, Bendix A, Labriola M, et al. Functional restoration versus outpatient physical training in chronic low back pain: a randomized comparative study. *Spine* 2000;25:2494–500.
19. Bronfort G, Goldsmith CH, Nelson CF, et al. Trunk exercise combined with spinal manipulative or NSAID therapy for chronic low back pain: a randomized, observer-blinded clinical trial. *J Manipulative Physiol Ther* 1996;19:570–82.
20. Buswell J. Low back pain: a comparison of two treatment programmes. *NZ J Physiother* 1982;10:13–7.
21. Calmels P, Jacob JF, Fayolle-Minon I, et al. Use of isokinetic techniques vs standard physiotherapy in patients with chronic low back pain: preliminary results [in French]. *Ann Readapt Med Phys* 2004;47:20–7.
22. Dalichau S, Scheele K. Effects of elastic lumbar belts on the effect of a muscle training program for patients with chronic back pain [in German]. *Z Orthop Ihre Grenzgeb* 2000;138:8–16.
23. Descarreaux M, Normand MC, Laurencelle L, et al. Evaluation of a specific home exercise program for low back pain. *J Manipulative Physiol Ther* 2002;25:497–503.
24. Deyo RA, Walsh NE, Martin DC, et al. A controlled trial of transcutaneous electrical nerve stimulation (TENS) and exercise for chronic low back pain. *N Engl J Med* 1990;322:1627–34.
25. Elnaggar IM, 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;16:967–72.
26. Frost H, Klaber Moffett JA, et al. Randomised controlled trial for evaluation of fitness programme for patients with chronic low back pain. *BMJ* 1995;310:151–4.
27. Frost H, Lamb SE, Doll HA, et al. Randomized controlled trial of physiotherapy compared with advice for low back pain. *Br Med J* 2004;329:708–11.
28. Gur A, Karakoc M, Cevik R, et al. Efficacy of low power laser therapy and exercise on pain and functions in chronic low back pain. *Lasers Surg Med* 2003;32:233–8.
29. Hansen FR, Bendix T, Skov P, et al. Intensive, dynamic back-muscle exercises, conventional physiotherapy, or placebo-control treatment of low back pain: a randomized, observer-blind trial. *Spine* 1993;18:98–107.
30. Hemmila HM, Keinanen-Kiukaanniemi SM, Levoska S. Does folk medicine work? A randomized clinical trial on patients with prolonged back pain. *Arch Phys Med Rehabil* 1997;78:571–7.
31. Johannsen F, Remvig L, Kryger P, et al. Exercises for chronic low back pain: a clinical trial. *J Orthop Sports Phys Ther* 1995;22:52–9.
32. Jousset N, Fanello S, Bontoux L, et al. Effects of functional restoration versus 3 hours per week physical therapy: a randomized controlled study. *Spine* 2004;29:487–93.
33. Kankaanpää M, Taimela S, Airaksinen O, et al. The efficacy of active rehabilitation in chronic low back pain: effect on pain intensity, self-experienced disability, and lumbar fatigability. *Spine* 1999;24:1034–42.
34. Kuukkanen TM, Malkia EA. An experimental controlled study on postural sway and therapeutic exercise in subjects with low back pain. *Clin Rehabil* 2000;14:192–202.
35. Manniche C, Hesselsoe G, Bentzen L, et al. Clinical trial of intensive muscle training for chronic low back pain. *Lancet* 1988;8626:1473–6.
36. Mannion AF, Muntener M, Taimela S, et al. A randomized clinical trial of three active therapies for chronic low back pain. *Spine* 1999;24:2435–48.
37. Moseley L. Combined physiotherapy and education is efficacious for chronic low back pain. *Aust J Physiother* 2002;48:297–302.
38. Niemisto L, Lahtinen-Suopanki T, Rissanen P, et al. A randomized trial of combined manipulation, stabilizing exercises, and physician consultation compared to physician consultation alone for chronic low back pain. *Spine* 2003;28:2185–91.
39. Petersen T, Kryger P, Ekdahl C, et al. The effect of McKenzie therapy as compared with that of intensive strengthening training for the treatment of patients with subacute or chronic low back pain: a randomized controlled trial. *Spine* 2002;27:1702–9.
40. Preyde M. Effectiveness of massage therapy for subacute low-back pain: a randomized controlled trial. *CMAJ* 2000;162:1815–20.
41. Rasmussen-Barr E, Nilsson-Wikmar L, Arvidsson I. Stabilizing training compared with manual treatment in sub-acute and chronic low-back pain. *Man Ther* 2003;8:233–41.
42. Risch SV, Norvell NK, Pollock M, et al. Lumbar strengthening in chronic low back pain patients: physiologic and psychological benefits. *Spine* 1993;18:232–8.
43. Rittweger J, Just K, Kautzsch K, et al. Treatment of chronic lower back pain with lumbar extension and whole-body vibration exercise: a randomized controlled trial. *Spine* 2002;27:1829–34.
44. Soukup MG, Glomsrod B, Lonn JH, et al. The effect of a Mensendieck exercise program as secondary prophylaxis for recurrent low back pain: a randomized, controlled trial with 12-month follow-up. *Spine* 1999;24:1585–92.
45. Lonn JH, Glomsrod B, Soukup MG, et al. Active back school. Prophylactic management for low back pain: a randomized, controlled, 1-year follow-up study. *Spine* 1999;24:865–71.
46. Torstensen TA, Ljunggren AE, Meen HD, et al. Efficiency and costs of medical exercise therapy, conventional physiotherapy, and self-exercise in patients with chronic low back pain: a pragmatic, randomized, single-blinded, controlled trial with 1-year follow-up. *Spine* 1998;23:2616–24.
47. Tritilanunt T, Wajanavisit W. The efficacy of an aerobic exercise and health education program for treatment of chronic low back pain. *J Med Assoc Thai* 2001;84(suppl 2):528–33.
48. Turner JA, Clancy S, McQuade KJ, et al. Effectiveness of behavioral therapy for chronic low back pain: a component analysis. *J Consult Clin Psychol* 1990;58:573–9.
49. Yelland MJ, Glasziou PP, Bogduk N, et al. Prolotherapy injections, saline injections, and exercises for chronic low-back pain: a randomized trial. *Spine* 2004;29:9–16.
50. Yeung CK, Leung MC, Chow DH. The use of electro-acupuncture in conjunction with exercise for the treatment of chronic low-back pain. *J Altern Complement Med* 2003;9:479–90.

51. Yozbatiran N, Yildirim Y, Parlak B. Effects of fitness and aquafitness exercises on physical fitness in patients with chronic low back pain. *Pain Clinic* 2004;16:35–42.
52. Zylbergold RS, Piper MC. Lumbar disc disease: comparative analysis of physical therapy treatments. *Arch Phys Med Rehabil* 1981;62:176–9.
53. Bentsen H, Lindgarde F, Manthorpe R. The effect of dynamic strength back exercise and/or a home training program in 57-year old women with chronic low back pain: results of a prospective randomized study with a 3-year follow-up period. *Spine* 1997;22:1494–500.
54. Galantino ML, Bzdewka TM, Eissler-Russo JL, et al. The impact of modified hatha yoga on chronic low back pain: a pilot study. *Altern Ther Health Med* 2004;10:56–9.
55. Hildebrandt VH, Proper KI, van den BR, et al. Cesar therapy is temporarily more effective in patients with chronic low back pain than the standard treatment by family practitioner: randomized, controlled and blinded clinical trial with 1 year follow-up [in Dutch]. *Ned Tijdschr Geneeskd* 2000;144:2258–64.
56. Kendall PH, Jenkins JM. Exercises for backache: a double-blind controlled trial. *Physiotherapy* 1968;54:154–7.
57. Lidstrom A, Zachrisson M. Physical therapy on low back pain and sciatica. *Scand J Rehabil Med* 1970;2:37–42.
58. Lie H, Frey S. Mobilizing or stabilizing exercise in degenerative disk disease in the lumbar region? [in Norwegian]. *Tidsskr Nor Laegeforen* 1999;119: 2051–3.
59. Pocock SJ, Hughes MD, Lee RJ. Statistical problems in the reporting of clinical trials: a survey of three medical journals. *N Engl J Med* 1987;317:426–32.
60. LeBlanc DC. *Statistics: Concepts and Applications for Science*. Sudbury, MA: Jones and Bartlett; 2004:155.