

Supplementary Material*

Chou R, Deyo R, Friedly J, Skelly A, Hashimoto R, Weimer M, Fu R, Dana T, Kraegel P, Griffin J, Grusing S. Nonpharmacologic therapies for low back pain: a systematic review for an American College of Physicians clinical practice guideline. *Ann Intern Med.* doi:10.7326/M16-2459

Supplement. Supplemental Tables

Supplement Table 1. Systematic Reviews of Nonpharmacologic Treatments for Low Back Pain

Supplement Table 2. Characteristics and Conclusions of Randomized Trials

Supplement Table 3. Quality Assessment of Systematic Reviews

Supplement Table 4. Quality Ratings of Randomized Trials

Supplement Table 5. Nonpharmacologic Treatments Versus Active Comparators for Acute or Subacute Low Back Pain

Supplement Table 6. Nonpharmacologic Treatments Versus Active Comparators for Chronic Low Back Pain

* This supplementary material was provided by the authors to give readers further details on their article. The material was reviewed but not copyedited.

Nonpharmacological Therapies for Low Back Pain: Systematic Review for an American College of Physicians Clinical Practice Guideline

Supplemental Tables

Supplement Table 1. Systematic reviews of nonpharmacologic treatments for low back pain

Author, Year	Number and Type of Studies	Interventions	Conclusions
<p>Acupuncture Lee, 2013 (58)</p>	<p>11 RCTs, Acute to subacute LBP (<12 weeks), 1139 patients (approximately 50 per arm), 5 low risk of bias</p>	<p>A. Acupuncture (n=3; 74 patients) B. Sham (n=3; 74 patients) C. Acupuncture (n=7; 500 patients) D. Conventional treatment (i.e., Meds) (n=7; 466 patients) E. Acupuncture + meds (n=1; 24 patients) F. Meds alone (n=1; 25 patients)</p>	<p>Moderate evidence of benefit in global improvement with acupuncture compared with NSAIDs, but the effect is very small. Inconsistent benefit of acupuncture compared with NSAIDs in terms of pain relief. Real acupuncture may be more effective than sham at reducing acute pain, but the effect is small and there appears to be no benefit in terms of function. Acupuncture in addition to medication appears more effective for pain relief and function than medication alone, but these differences are small.</p>
<p>Lam, 2013 (59)</p>	<p>32 studies, 25 in meta-analysis (n=6266 patients); 7 low risk of bias, Duration of LBP: 4 trials Subacute to chronic LBP (>6 weeks), 28 trials chronic (>3 months), duration of followup 0-48 months</p>	<p>A. Acupuncture (n=5; 1735 patients) B. No treatment (n=5; 1596 patients) C. Acupuncture (n=3; 75 patients) D. Medication (n=3; 80 patients) E. Acupuncture (n=3; 68 patients) F. TENS, (n=3; 72 patients) G. Acupuncture (n=4; 447 patients) H. Sham (n=4; 452 patients) acupuncture, I. Acupuncture in addition to usual care (n=4; 139 patients) J. Self-care or usual care, (n=4; 139 patients) K. Electroacupuncture (n=6; 156 patients) L. Usual care.(n=6; 162 patients)</p>	<p>Acupuncture improved pain and function immediately post intervention more than no treatment, sham acupuncture or medications such as NSAIDs, muscle relaxants or analgesics, but these differences were small. Patients who received acupuncture in addition to usual care had greater pain relief and improved function immediately postintervention and at followup compared with those who received usual care alone. Patients who received electroacupuncture reported significantly less pain and levels of activity limitation than the control group immediately postintervention and at followup. There was no evidence that acupuncture was better than TENS.</p>

Author, Year Exercise	Number and Type of Studies	Interventions	Conclusions
Bystrom, 2013 (13)	16 RCTs (1 with 2 arms) (n=1933) 80% with CBLP; included studies of subacute if duration >6 months; define sub acute as 4-12 weeks short (6 weeks–4 months), intermediate (4–8 months) and long term (8-15 months) followup	A. Motor Control Exercises (MCE) B. General exercise (n=7 [1 with two arms]; 741 patients) C. Minimal intervention (n=3; 541 patients) D. Multimodal physical therapy (n=4; 499 patients) E. MCE as part of multimodal intervention versus other components of that intervention (n=2; 152 patients)	For chronic low back pain, MCE was associated with lower pain intensity versus general exercise: Short term (6 trials, WMD -7.80 on 0 to 100 scale, 95% CI -10.95 to -4.65) Intermediate term (3 trials, WMD -6.06, 95% CI -10.94 to -1.18) Effects were smaller and not statistically significant at long term (4 trials, WMD -3.10, 95% CI -7.03 to 0.83) MCE was also associated with better function: Short term (6 trials, WMD -4.65 on 0 to 100 scale, 95% CI -6.20 to -3.11) Long term (3 trials, WMD -4.72, 95% CI -8.81 to -0.63). For chronic low back pain, MCE was associated with lower pain scores versus minimal intervention: Short term (WMD -12.48 on a 0 to 100 scale, 95% CI -19.04 to -5.93) Intermediate term (WMD -10.18, 95% CI -16.64 to -3.72) Long term (WMD -13.32 95% CI -19.75 to -6.90) MCE was also associated with better function: Short term (3 trials WMD -9.00 on 0 to 100 scale, 95% CI -15.28 to -2.73) Intermediate term (2 trials WMD -5.62, 95% CI -10.46 to -0.77) Long term (2 trials, WMD -6.64, 95% CI -11.72 to -1.57)

Author, Year	Number and Type of Studies	Interventions	Conclusions
Oesch, 2010 (15)	23 RCTs (n=4138) (20 with data for meta-analysis, 17 comparisons of exercise vs. usual care and 11 comparisons of two different exercise)	A. Exercise (n=23) B. Usual care (n=17)	No effects on work disability at short-term (~4 weeks) or intermediate-term (~6 months) followup, based on pooled analyses of high-quality studies (6 comparisons in 5 trials, OR 0.80, 95% CI 0.51 to 1.25 and 5 comparisons in 4 trials, OR 0.78, 95% CI 0.45 to 1.34, respectively). Exercise was associated with lower likelihood of work disability at long-term (~12 months) followup (10 comparisons in 8 trials, OR 0.66, 95% CI 0.48 to 0.92).
van Middelkoop, 2010 (14)	nonacute nonspecific LBP, duration ≥ weeks 37 RCTs (N=3957) chronic (≥12 weeks) nonspecific LBP post-treatment, short, intermediate, and long-term followup (not defined)	A. Exercise B. Wait list/no treatment (8 trials) C. Usual care (6 trials) D. Back school/education (3 trials) E. Other forms of exercise therapy (11 trials)	Exercise therapy was associated with decreased pain intensity (3 trials, WMD -9.23, 95% CI -16.02 to -2.43) and better function (3 RCTs, WMD -12.35 on a 0 to 100 scale, 95% CI -23.0 to -1.69) versus usual care at the end of treatment. Effects on function were smaller but remained statistically significant at intermediate- and long-term followup (mean differences -5.23 and -3.17). Effects on pain were also smaller, and no longer statistically significant at long-term followup (mean difference -4.94, 95% CI -10.45 to 0.58).
Massage Furlan, 2010 (111)	13 RCTs (n=39 to 262, total 1596) Duration of followup: Immediately after sessions to 52 weeks (42 weeks after completion of therapy) Duration of low back pain: acute (1 RCT), subacute to chronic (4 RCTs), chronic (8 RCTs)	A. Massage (111 patients) B. Sham/placebo massage (n= 2 RCTs, 111 patients) C. Massage (1026 patients) D. Other treatments (manipulation [1 RCT, 67 patients]), exercise [1 RCT, 47 patients]), relaxation therapy [3 RCTs, 297 patients]), acupuncture [1 RCT, 172 patients]), physiotherapy [2 RCTs, 275 patients]), self-care education [1 RCT, 168 patients]) E. Massage + other intervention F. Other intervention (exercise and education [1 RCT, 47 patients], exercise [2 RCTs, 290 patients], usual care [2 RCTs, 183 patients]) without massage G. Swedish massage H. Acupuncture massage (1 RCT, 190 patients) or traditional Thai massage (1 RCT, 180 patients)	Moderate evidence of short and long (up to 1 year) term improvement of pain and function with massage as compared with sham/placebo or other treatments, but the differences in improvement are small. Massage appears to be most beneficial when added to exercise and/or education. One RCT suggests acupuncture massage is superior to Swedish massage, otherwise there appears to be no difference between massage techniques, although evidence is limited.

Multidisciplinary Rehabilitation

Author, Year	Number and Type of Studies	Interventions	Conclusions
Kamper, 2014 (54)	41 RCTs; all chronic low back pain; Multidisciplinary biopsychosocial rehab (MBR) vs. usual care: 16 trials; MBR vs. physical treatment: 19 trials; MBR vs. waitlist: 4 trials	<ol style="list-style-type: none"> 1. MBR vs. usual care <ol style="list-style-type: none"> a. Short-term pain outcomes, 9 trials, 879 patients) b. Long-term pain outcomes, 7 trials, 821 patients) c. Short-term disability outcomes, 9 trials, 939 patients. d. Long-term disability outcomes, 6 trials, 722 patients e. Short-term work outcomes, 2 trials, 373 patients f. Long-term work outcomes, 7 trials, 1360 patients 2. MBR vs. physical treatment <ol style="list-style-type: none"> a. Short-term pain outcomes, 12 trials, 1661 patients b. Long-term pain outcomes, 9 trials, 872 patients c. Short-term disability outcomes, 13 trials, 1878 patients d. Long-term disability outcomes 10 trials, 1169 patients e. Short-term work outcomes, 3 trials, 379 patients f. Long-term work outcomes 8 trials, 1006 patients 3. MBR vs. waitlist <ol style="list-style-type: none"> a. Short-term pain outcomes, 3 trials, 213 patients b. Short-term disability outcomes, 3 trials, 213 patients 	There is evidence that MBR improves pain and disability more than usual care in the short and long term, but no evidence that it improves work outcomes in the short or long term. There is evidence that MBR improves pain and disability more than no MBR in the short term. There is evidence that MBR improves pain, disability, and work outcomes more than physical treatments in the short and long term.

Psychological Therapies

Author, Year	Number and Type of Studies	Interventions	Conclusions
Henschke, 2010 (48)	28 RCTs Chronic LBP: 28 trials Subacute, acute LBP: 0 trials Psychological therapy vs. waiting list: 12 trials Psychological therapy vs. other noninvasive interventions: 7 trials One psychological therapy vs. another: 10 trials Psychological therapy plus other intervention vs. other intervention alone: 9 trials	<p>A. Psychological therapy vs. waiting list (12 trials total)</p> <ol style="list-style-type: none"> 1. Respondent therapy (relaxation training) vs. wait list: n=74 (3 trials) 2. Respondent therapy (EMG biofeedback) vs. wait list: n=108 (4 trials) 3. Operant therapy vs. wait list: n=243 (4 trials) 4. Cognitive therapy vs. wait list: n=68 (2 trials) 5. Combined psychological therapies (including CBT) vs. wait list: n=239 (5 trials) <p>B. Psychological therapy vs. other intervention (7 trials total):</p> <ol style="list-style-type: none"> 1. Psychological therapy vs. usual care: n=330 (2 trials) 2. Psychological therapy vs. group exercise : n=146 (2 trials) 3. Psychological therapy vs. guideline-based care: n=114 (1 trial) 4. Psychological therapy vs. Back education: n=36 (1 trial) 5. Psychological therapy vs. hypnosis: n=15 (1 trial) <p>C. One psychological therapy vs. another (10 trials total):</p> <ol style="list-style-type: none"> 1. Respondent (EMG biofeedback) vs. respondent (relaxation therapy) therapy: n=24 (1 trial) 2. Cognitive vs. operant therapy: n=93 (2 trials) 3. Cognitive vs. respondent therapy: n=93 (1 trial) 4. Combined psychological therapies vs. cognitive therapy: n=61 (2 trials) 5. Combined psychological therapies vs. operant therapy: n=278 (4 trials) 6. Combined psychological therapies vs. respondent therapy: n=97 (4 trials) <p>D. Psychological therapy plus other intervention vs. other intervention alone (9 trials total):</p> <ol style="list-style-type: none"> 1. Physiotherapy with or without psychological therapy: n=59 (2 trials) 2. Exercise with or without psychological therapy: n=262 (3 trials) 3. Inpatient rehabilitation with or without psychological therapy: n=435 (3 trials) 4. Education booklet/audio cassette with or without psychological therapy: n=234 (1 trial) <p>N=3090 total</p>	Moderate evidence of post-treatment pain relief benefit with operant therapy versus waiting list, and with psychological therapy versus usual care. Moderate evidence that there is no benefit of one type of psychological therapy over another in pain relief through six months. Moderate evidence of no benefit of psychological therapy over group exercise for pain relief or depression through twelve months. Otherwise, there was only low or very low evidence available for other comparisons and/or outcomes. All conclusions are for the chronic low back pain patient population.

Spinal Manipulation

Author, Year	Number and Type of Studies	Interventions	Conclusions
Rubinstein, 2012 (86)	20 RCTs: 9 acute LBP; 4 mixed acute and subacute LBP; 6 any LBP Duration of followup <3 months to > 12 months. More than half of the studies limited followup to short-term measurements only (that is < 3 months) including, in particular, one study that measured the effect two days post-treatment only (Sutlive 2009). Five studies measured the long-term (that is > 12 months) effects of the treatments.	A. Any SMT (n=20) 1. Thrust SMT (n=13) 2. Combination mobilization, manipulation or both SMT (n=4) or unclear (n=3) B. Other active interventions (exercise; physical therapy; massage; standard care; back school; n=8) C. Sham SMT (n=1) D. Inert interventions (education; ultrasound alone; ultrasound + cold; ultrasound; short-wave diathermy; anti-edema gel; bed rest; n=7)	Low to very low-quality evidence of no difference in effect of SMT compared with inert interventions, sham SMT, or when added to another intervention, in terms of pain, function, QOL, work, global improvement. Low to mod no diff vs. other interventions, with the exception of moderate short-term effect of SMT on functional status when added to another intervention.
Rubinstein, 2011 (87)	26 total studies with wide variety of comparisons, 9 with low risk of bias, LBP >12 weeks, 18+ years old, outcomes short, intermediate and long term (>12 months)	A. Any SMT (n=26) B. Inert interventions (i.e., detuned short-wave diathermy and detuned ultrasound; n=4) C. Other active interventions (exercise; physical therapy; massage; standard care; back school; n=15) D. Sham SMT (n=3)	SMT has statistically significant short-term effect on pain and function compared with other interventions; varying quality that SMT has a statistically significant short-term effect on pain and function when SMT is added to another intervention. Effect sizes were small - not clinically relevant. Very low-quality evidence that SMT is no more effective than inert interventions or sham SMT for short-term pain relief or functional status.

Yoga

Author, Year	Number and Type of Studies	Interventions	Conclusions
Cramer, 2013 (37)	<p>10 RCTs in qualitative synthesis;</p> <p>Two citations with different outcomes from same trial, treated as single study</p> <p>8 included in quantitative synthesis;</p> <p>9/10 studies included CLBP patients; 1 included acute, subacute or chronic</p>	<p>A. Yoga B. Usual care C. Education D. Exercise</p> <p>TOTAL n for each intervention unclear across all studies;</p> <p>Total N for all studies=1067</p>	<p>For chronic low back pain, yoga was associated with lower pain intensity and better function versus exercise in most trials, though effects were small and differences were not always statistically significant</p> <p>For chronic low back pain, yoga was associated with: Lower short-term pain intensity versus education (5 trials, SMD -0.45, - 95% CI -0.63 to -0.26; I²=0%) Effects were smaller and not statistically significant at longer-term followup (4 trials, SMD -0.28, 95% CI-0.58 to -0.02' I²=47%) Yoga was also associated with better function: Short-term (5 trials, SMD 0.45, 95% CI-0.65 to -0.25; I²=8%) Long term followup (4 trials, SMD 0.39, 95% CI -0.66 to -0.11; I²=40%)</p>

CLBP=chronic low back pain; CI=confidence interval; EMG=electromyography; LBP=low back pain; MBR=multidisciplinary biopsychosocial rehabilitation; MCE=motor control exercises; NSAID=Nonsteroidal anti-inflammatory drug; OR=odds ratio; QOL=quality of life; RCT=randomized controlled trial; SMT=standard mean difference; TENS=transcutaneous electrical nerve stimulation; WMD=weighted mean difference

Supplement Table 2. Characteristics and conclusions of randomized trials

Author, Year Duration of Followup LBP Duration <i>Quality</i>	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Exercise Trials				
Albaladejo, 2010 (16) Subacute, chronic <i>Fair</i>	A. Education + 4 sessions of physiotherapy (n=100) B. Education (n=139) C. Usual care (n=109)	A. vs. B. vs. C. Median age: 51 vs. 51 vs. 53 Female: 68% vs. 63% vs. 72% Median pain intensity: 7.5 vs. 8 vs. 8 Median RDQ: 9.5 vs. 9.0 vs. 7.5 Median CSQ: 7.0 vs. 8.0 vs. 6.0 Median SF-12 PCS: 34.8 vs. 35.8 vs. 36.5 Median SF-12 MCS: 44.6 vs. 50.1 vs. 49.8	A. vs. B. vs. C. Change in median VAS (0- 10), low back pain: -2.0 vs. -2.0 vs. 0 Change in median VAS (0- 10), referred pain: -2.0 vs. - 2.0 vs. -0.5	A. vs. B. vs. C. Improvement in RDQ: 2.0 vs. 1.6 vs. -0.3 Change in CSQ: -1.0 vs. -1.0 vs. 2.0 Change in SF-12 PCS: -3.2 vs. -2.4 vs. 0.6 Change in SF-12 MCS: -2.8 vs. -1.8 vs. 6.1

Author, Year
Duration of
Followup
LBP Duration
Quality

Intervention and Duration of Treatment

Population

Pain Outcomes

Other Outcomes

Albert, 2012 (17) 12 months Acute, subacute, chronic <i>Fair</i>	<p>A: Symptom-guided exercises (n=95). Directional end-range exercises and postural instructions guided by the individual patient's directional preference (based on the McKenzie method); stabilizing exercises for the transverse abdominis and multifidus muscles and dynamic exercises for the outer layers of the abdominal wall and back extensors; all patients received home exercise programs</p> <p>B: Sham exercises (n=96). Optional exercises that were not back related but were low-dose exercises to simulate an increase in systemic blood circulation.</p> <p>Both groups received identical information and advice and optional paracetamol and/or NSAIDs. Treatment lasted for 8 weeks with a minimum of 4 and a maximum of 8 treatments. Patients were discouraged from receiving any additional treatment of their sciatica.</p>	<p>A. vs. B. Mean age: 46 vs. 44 Female: 43% vs. 53% Baseline Current leg pain (LBPRS): 4.3 ± 2.3 vs. 4.5 ± 2.5 Total leg pain, median (IQR): 18 (15–21) vs. 18 (12–21); p=NS Disability (RDQ), median (IQR): 16 (11–18) vs. 15 (12–18) Quality of Life: 0.62 ± 0.18 vs. 0.62 ± 0.62</p>	<p>A. vs. B. Current leg pain (LBPRS) (mean, SD) 8 weeks (end of treatment): 1.5 ± 2.1 vs. 2.3 ± 2.7; p=0.06 EPC calculation of test mean diff -0.8 (95% CI -0.09 to -1.15) 12 months: 1.5 ± 2.1 vs. 1.4 ± 2.4; p=NS Total leg pain (LBPRS) (median, IQR) 8 weeks: 4 (0–9) vs. 4 (0–12); p=NS 12 months: 3 (0–10) vs. 2 (0–8); p=NS</p>	<p>A. vs. B. Disability (RDQ) (median, IQR) 8 weeks: 6 (2–12) vs. 6 (2–12); p=NS 12 months: 3.5 (1–10) vs. 3.5 (1–10); p=NS ≥30% improvement from baseline: 73% vs. 77.5%; p=NS Quality of Life (EQ-5D) (mean, SD) 12 months: 0.82 ± 0.21 vs. 0.79 ± 0.24; p=NS Global improvement 8 weeks Much better: 80% vs. 60% Some better: 14% vs. 26% 12 months: Much better: 84% vs. 76% Some better: 16% vs. 18% Group A significantly (p<0.008) more improved (better or much better) compared with group B at both time points Patient satisfaction: 93.5% vs. 90.5%; p=NS</p>
Bronfort, 2011 (18) 52 weeks Chronic <i>Good</i>	<p>A. Supervised exercise therapy for 12 weeks (n=100)</p> <p>B. Chiropractic spinal manipulation for 12 weeks (n=100)</p> <p>C. Home exercise and advice for 12 weeks (n=101)</p>	<p>A. vs. B. vs. C. Mean age: 44.5 vs. 45.2 vs. 45.6 years Female: 57% vs. 66% vs. 58% Mean pain severity score (0-10): 5.1 vs. 5.4 vs. 5.2 Roland-Morris disability score (0-23): 8.4 vs. 8.7 vs. 8.7</p>		<p>Only significant between-group differences in patient-reported outcomes were for satisfaction (favoring A, p<0.01 at 12 weeks and p<0.001 at 52 weeks) Overall treatment effect was significant for endurance (p<0.05) and strength (p<0.05) but not range of motion (also favoring A).</p>

**Author, Year
Duration of
Followup
LBP Duration
Quality**

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Garcia, 2013 (19) 1, 3, 6 months Acute, subacute, chronic Good	<p>A: McKenzie method (n=74). Exercises and progression tailored to the individual. Included a basic educational component and guidance on completing the exercises at home. Patients with a direction preference for extension were instructed to use a back roll while sitting.</p> <p>B: Back school (n=74). New exercises were prescribed and progressed following the sequence proposed by the program (i.e., not tailor to the individual). Educational component and theoretical and practical information given. All sessions except for the first were conducted in a group setting.</p> <p>All patients received 4 one-hour sessions over 4 weeks. In all patients, directional preference was assessed at baseline and the treating therapist was informed before the randomization. All patients received information in order to maintain lordosis while sitting without exacerbating their symptoms</p>	<p>A. vs. B. Mean age: 53.7 vs. 54.2 years Female: 78.4% vs. 68.9% Duration of LBP: 21 vs. 24 months Recent episode of LBP: 62.2% vs. 63.5% Pain intensity (NRS, 0-10): 6.77 ± 2.12 vs. 6.41 ± 2.54 Disability (RDQ, 0-24): 11.32 ± 4.95 vs. 11.08 ± 5.84 Quality of life (WHOQOL-BREF, 0-100) Physical domain: 51.64 ± 14.49 vs. 51.49 ± 17.05 Psychological domain: 62.88 ± 15.86 vs. 60.11 ± 15.86 Social domain: 63.62 ± 18.27 vs. 63.15 ± 18.96 Environmental domain: 55.40 ± 13.66 vs. 54.74 ± 16.09</p>	<p>A. vs. B. Unadjusted mean difference ± SD for A. vs. B.; adjusted mean difference (95% CI) for B – A Pain intensity (NRS, 0–10) 1 month: 4.14 ± 2.87 vs. 4.39 ± 2.73; 0.66 (–0.29 to 1.62), p=0.17 3 months: 5.18 ± 2.61 vs. 5.53 ± 2.78; 0.71 (–0.23 to 1.67), p=0.14 6 months: 5.09 ± 2.89 vs. 5.19 ± 3.08; 0.48 (–0.47 to 1.43), p=0.32</p>	<p>A. vs. B. Unadjusted mean difference ± SD for A. vs. B.; adjusted mean difference (95% CI) for B – A Disability (RDQ, 0–24) 1 month: 6.20 ± 5.06 vs. 8.15 ± 5.79; 2.37 (0.76 to 3.99), p=0.004 3 months: 7.12 ± 5.67 vs. 8.39 ± 6.30; 1.51 (–0.09 to 3.11), p=0.06 6 months: 6.77 ± 6.02 vs. 8.12 ± 6.45; 1.55 (–0.05 to 3.16), p=0.06 Achievement of ≥5-point improvement: 53% (39/74) vs. 30% (22/73), p=0.01; RR 1.8, 95% CI 1.2 to 2.7 Quality of Life (WHOQOL-BREF, 0-100) Physical domain 1 month: 62.45 ± 16.94 vs. 59.27 ± 16.88; –3.65 (–8.26 to 0.96), p=0.12 3 months: 62.25 ± 15.37 vs. 57.43 ± 17.76; –4.67 (–9.26 to –0.07), p=0.04 6 months: 61.48 ± 16.12 vs. 60.76 ± 18.87; –0.44 (–5.04 to 4.16), p=0.85 Psychological domain 1 month: 67.68 ± 15.15 vs. 65.12 ± 13.98; –0.18 (–4.17 to 3.80), p=0.92 3 months: 67.62 ± 16.07 vs. 65.14 ± 14.14; 0.14 (–3.82 to 4.11), p=0.94 6 months: 68.00 ± 14.18 vs. 66.72 ± 14.15; 1.50 (–2.48 to 5.47), p=0.46 Social domain 1 month: 67.45 ± 18.00 vs. 67.24 ± 15.96; –0.47 (–5.50 to 4.56), p=0.85 3 months: 69.03 ± 16.11 vs. 65.76 ± 16.00; –3.15 (–8.16 to 1.85), p=0.21 6 months: 66.00 ± 18.74 vs. 66.09 ± 15.00; 0.26 (–4.75 to 5.28), p=0.91 Environmental domain 1 month: 58.57 ± 14.82 vs. 57.62 ± 16.48; –0.51 (–4.06 to 3.03), p=0.77 3 months: 58.23 ± 14.65 vs. 56.16 ± 14.75; –1.41 (–4.94 to 2.12), p=0.43 6 months: 57.84 ± 14.61 vs. 57.44 ± 15.00; 0.29 (–3.24 to 3.83), p=0.87</p>

*RR (95% CI) calculated by EPC

Author, Year
Duration of
Followup
LBP Duration
Quality

	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
George, 2008 (20) 6 months Acute, subacute Poor	<p>A: Treatment-based classification + Graded exposure (GX) (n=33). Fearful activities assessed; top 2 most feared activities implemented under this protocol using progression based on 0-10 NRS fear rating and performed under supervision of physical therapy and clinical staff. Also received patient education materials focused on biopsychosocial model.</p> <p>B: Treatment-based classification + Graded activity (GA) (n=35). Parameters (duration, intensity, and frequency) used to reach pain tolerance were then established as the activity quota; graded activity principles were used to progress exercise during subsequent treatment sessions. Also received patient education materials focused on biopsychosocial model</p> <p>C: Physical therapy based on the treatment-based classification (TBC) system (Delitto et al.) (n=34). Also received educational materials that were anatomically focused.</p>	<p>A. vs. B. vs. C. Mean age: 40.1 vs. 37.6 vs. 34.9 years Female: 64% vs. 69% vs. 68% Baseline Pain (NRS, 0-10): 4.7 ± 2.1 vs. 5.2 ± 1.8 vs. 4.3 ± 2.0 Function (PIS): 3.1 ± 1.6 vs. 3.6 ± 2.1 vs. 2.9 ± 1.7 Disability (ODI): 30.7 ± 15.6 vs. 31.1 ± 15.8 vs. 29.2 ± 15.7</p>	<p>A. vs. B. vs. C. Pain intensity (NRS, 0–10) <u>High fear</u> Baseline: 5.1 ± 2.1 vs. 5.1 ± 1.9 vs. 5.1 ± 1.8 4 weeks: 2.1 ± 2.0 vs. 2.3 ± 2.1 vs. 2.0 ± 1.6 6 months: 2.1 ± 2.3 vs. 1.5 ± 2.1 vs. 1.6 ± 1.3 <u>Low fear</u> Baseline: 3.9 ± 1.5 vs. 4.9 ± 2.1 vs. 3.1 ± 2.1 4 weeks: 1.7 ± 0.9 vs. 2.1 ± 2.1 vs. 1.8 ± 1.9 6 months: 1.0 ± 1.0 vs. 2.3 ± 1.7 vs. 1.0 ± 1.2</p> <p>Effect sizes Pain intensity (NRS, 0-10) 4 weeks A. vs. B.: 0.11 A. vs. C.: -0.05 B vs. C.: -0.16 6 months A. vs. B.: -0.32 A. vs. C.: -0.26 B vs. C.: 0.01</p> <p>p=NS for all comparisons. These post hoc effect sizes suggest that for the primary comparisons of interest (GX vs. GA and GX vs. TBC) total sample sizes needed to detect these magnitudes of differences would range from 114 to over 700.</p>	<p>A. vs. B. vs. C. Disability (ODI, 0–100) <u>High fear</u> Baseline: 32.3 ± 16.3 vs. 29.9 ± 18.4 vs. 32.9 ± 16.1 4 weeks: 16.5 ± 12.1 vs. 11.5 ± 11.8 vs. 16.4 ± 14.9 6 months: 16.7 ± 17.6 vs. 11.3 ± 14.2 vs. 11.4 ± 11.5 <u>Low fear</u> Baseline: 20.4 ± 13.1 vs. 30.4 ± 13.3 vs. 23.0 ± 15.5 4 weeks: 11.4 ± 11.6 vs. 16.7 ± 11.9 vs. 12.0 ± 11.5 6 months: 9.7 ± 8.2 vs. 15.8 ± 11.1 vs. 5.8 ± 7.1</p> <p>Effect sizes Disability (ODI, 0-100) 4 weeks A. vs. B.: -0.40 A. vs. C.: -0.02 B vs. C.: 0.39 6 months A. vs. B.: -0.38 A. vs. C.: -0.37 B vs. C.: 0.01</p> <p>p=NS for all comparisons. These post hoc effect sizes suggest that for the primary comparisons of interest (GX vs. GA and GX vs. TBC) total sample sizes needed to detect these magnitudes of differences would range from 114 to over 700.</p> <p>Proportion of Success vs. Failure (ODI >10 point change, NRS >2 point change) at 6 months NRS 46% vs. 43% vs. 41% ODI 43% 41%, 56% p=0.70</p>

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Hagen, 2010 (21) 24 months LBP duration not reported <i>Fair</i>	A: Standardized physical exercise program (n=124). Aim was to re-educate the trunk muscle to its normal stabilizing role and to improve balance, muscle coordination, and proprioception; program included warm-up (8 minutes), circuit training (34 minutes), stretching (13 minutes), and relaxation (5 minutes); duration 1 hour, 3x/week for 8 weeks. B: No treatment (n=122). Received a brief intervention program before randomization.	A. vs. B. Mean age: 40.7 vs. 41.6 years Female: 52% vs. 50%	A. vs. B. No statistically significant difference between groups at any followup time point - 6, 12, 18 or 24 months – for Pain intensity.	A. vs. B. Only statistically significant difference found was for the sock test (physical function), which was more improved in Group A. vs. B.: mean difference -0.34; 95% CI -0.66 to -0.01; p=0.041 (time point NR). No statistically significant difference between groups at any followup time point - 6, 12, 18 or 24 months - for the following (no data provided): Functional tests (pick-up test, loaded reach test, 15 meter walk, fingertip-to-floor test, static balance test) Physical activity Walking distance Disability (RDQ) Subjective health complaints Psychological distress (HSCL-25) Return to work
Hartvigsen, 2010 (22) 52 weeks Acute, subacute, chronic <i>Fair</i>	A. Supervised Nordic walking in groups twice/week for 8 weeks (n=45) B. Nordic walking instruction for 1 hour, with instruction to continue independently (n=46) C. Active living and exercise information (n=45)	A. vs. B. vs. C. Mean age: 49.2 vs. 45.4 vs. 45.5 years Female: 76% vs. 69% vs. 68% LBP rating scale (0-100), pain: 46.1 vs. 50.7 vs. 47.3 LBP rating scale (0-100), function: 44.4 vs. 47.3 vs. 48.9 Patient-specific function scale (0-100): 18.4 vs. 20.1 vs. 17.3 EQ-5D (0-100): 67.5 vs. 62.7 vs. 63.9	A. vs. B. vs. C. Mean improvement at 8 weeks in LBP rating scale, pain: 8.8 vs. 3.4 vs. 4.8; significant at all time-points for group A, significant only at 8 and 26 weeks for group B, significant only at 8 weeks for group C; no significant between-group differences at any point	A. vs. B. vs. C. Mean improvement at 8 weeks in LBP rating scale, function: 7.4 vs. 3.2 vs. 3.8; significant at all time-points for group A, never significant for group B, and significant only at 8 and 26 weeks in group C; no significant between-group differences at any point Patient-specific function scale: all groups improved significantly from baseline, but there were no between-group differences EQ-5D: very small and similar changes in all groups

Author, Year
Duration of
Followup
LBP Duration
Quality

Author, Year	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Helmhout, 2008 (23) Acute, subacute, chronic Poor	<p>A: Lumbar extensor strength training program (n=71). Standardized, progressive resistance training of the isolated lumbar extensor muscle groups aimed at both strength and endurance gain; duration 10 weeks, 14 sessions 2x/wk and 3 isometric back strength tests (in weeks 1, 5, and 10). Training sessions were carried out on a Total Trunk Rehab machine. Patients were not allowed to undergo cotreatments during the treatment period.</p> <p>B: Regular physical therapy program (n=56). Regular physical therapy for 10 weeks, or less when the patient was free of complaints; could include hands-on treatment (e.g., passive mobilizing and pain cushioning techniques, manual therapy) and/or hands-off treatment (e.g., exercise therapy, individual education, instruction on the back function) (in the Dutch army, active therapy forms are favored); no cotreatments allowed, nor exercise on equipment that mimicked the specific components of the lower back machine</p>	<p>A. vs. B. Mean age: 37 vs. 35 years Female: 3% vs. 4% Baseline Function (PSFS): 178 ± 65 vs. 178 ± 52 Disability (RDQ): 8.3 ± 4.8 vs. 7.9 ± 4.4 Back extension strength (net muscular torque): 214 ± 64 vs. 212 ± 65</p>	<p>A. vs. B. (mean ± SD; between group difference, 95% CI) LBP episodes 6 months (back pain in 1st half of year after the end of the treatment period?) (A, n=56; B, n=40): No, not at all: 9% vs. 18% Yes, incidentally: 57% vs. 63% Yes, monthly: 11% vs. 3% Yes, weekly: 23% vs. 18% 12 months (back pain in 2nd half of year after the end of the treatment period?) (A, n=61; B, n=46): No, not at all: 25% vs. 22% Yes, incidentally: 55% vs. 50% Yes, monthly: 2% vs. 11% Yes, weekly: 18% vs. 17%</p>	<p>A. vs. B. (mean ± SD; between group difference, 95% CI) Function (PSFS, score 0–300) 5 weeks: 119 ± 70 (n=64) vs. 116 ± 67 (n=46) 10 weeks: 85 ± 72 (n=59) vs. 97 ± 74 (n=47); –0.608 (–2.693 to 1.477), p=0.57 36 weeks: 74 ± 72 (n=57) vs. 64 ± 59 (n=37) 62 weeks: 69 ± 71 (n=61) vs. 65 ± 69 (n=45); –0.136 (–0.344 to 0.616), p=0.58 Disability (RDQ, score 0–24) 5 weeks: 5.8 ± 4.8 (n=64) vs. 4.2 ± 4.2 (n=46) 10 weeks: 3.4 ± 4.6 (n=59) vs. 3.5 ± 4.2 (n=47); –0.025 (–0.134 to 0.085), p=0.66 36 weeks: 3.2 ± 4.3 (n=57) vs. 2.7 ± 3.8 (n=37) 62 weeks: 2.6 ± 4.4 (n=61) vs. 2.5 ± 3.9 (n=45); 0.000 (–0.025 to 0.026), p=0.99 Global perceived effect (GPE) 5 weeks: no data 10 weeks: 2.4 ± 0.8 (n=59) vs. 2.4 ± 0.7 (n=47) 36 weeks: 2.5 ± 1.0 (n=57) vs. 2.3 ± 0.9 (n=37) 62 weeks: 2.2 ± 1.0 (n=61) vs. 2.3 ± 1.0 (n=45); –0.002 (–0.010 to 0.006), p=0.66 Patient satisfaction (very satisfied; final degree of satisfaction at end of treatment program): 89% (n=56) vs. 89% (n=46) Back extension strength (net muscular torque) 5 weeks: 23 ± 62 (n=64) vs. 246 ± 74 (n=46) 10 weeks: 244 ± 66 (n=59) vs. 247 ± 73 (n=47) 36 weeks: 264 ± 64 (n=57) vs. 254 ± 73 (n=37) 62 weeks: 267 ± 62 (n=61) vs. 249 ± 74 (n=45) p=NS for all time points</p>

Author, Year
Duration of
Followup
LBP Duration
Quality

Author, Year	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Henchoz, 2010 (24) Subacute, chronic Poor	A. Functional multidisciplinary rehabilitation, followed by a 12-week exercise program (n=56) B. Functional multidisciplinary rehabilitation, followed by usual care (n=49)	A. vs. B. Mean age: 41 vs. 39 years Female: 34% vs. 45% Mean VAS (0-100): 5.3 vs. 5.1	A. vs. B. VAS (0-100): 3.8-3.8 (p=0.521) vs. 3.6-3.8 (p=0.995)	A. vs. B. ODI: 30.2-25.3 (p<0.001) vs. 30.5-27.2 (p=0.059) SFS: 66.1-89.8 (p<0.05) vs. 65.5-78.8 (p=0.653) Sorensen test (s): 64.8-81.6 (p<0.05) vs. 67.1- 63.9 (p=0.249) MMS test, flexion (cm): 5.65-5.15 (p=0.368) vs. 5.27-5.19 (p=0.561) MMS test, extension (cm): -1.63 to -1.61 (p=0.138) vs. -1.46 to -1.64 (p=0.353) Fingertip-floor distance (cm): 126.5-135.7 (p=0.076) vs. 129.1-136.0 (p=0.470) Shirado test (s): 11.3-8.0 (p=0.063) vs. 17.3- 10.0 (p<0.001) Modified Bruce test (min): 11.2-8.4 (p<0.001) vs. 11.2-8.7 (p<0.001)

Author, Year
Duration of
Followup
LBP Duration
Quality

	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Hofstee, 2002 (25) 6 months Acute Poor	<p>A: Physiotherapy (n=83). The protocol consisted of instructions and advice, segmental mobilization, disc unloading and loading exercises, depending on patients' conditions, and hydrotherapy; 2x/week for at least 4 to 8 weeks; asked to perform daily exercises at home.</p> <p>B: Bed rest (at home or in hospital) (n=84). Instructed to stay in bed for 7 days; only allowed out of bed to use the bathroom and shower. After this period, patients supposed to rest as much as possible when in pain.</p> <p>C: Continuation of activities of daily living (control group) (n=83). Continue jobs, household activities, studies, or hobbies to the best of the patients' abilities; advised to adjust the intensity, duration, and frequency of their activities according to the pain they experienced.</p> <p>All patients received a brochure with instructions and advice regarding their respective treatment; were allowed to use analgesic medication and to call the investigator for help if they had problems or questions. When patients called, they were reassured and urged to comply with their assigned treatment; if necessary, they were seen at the outpatient clinic.</p>	<p>A. vs. B. vs. C. Mean age: 38 vs. 38 vs. 41.9 years; p=0.02 Female: 37% vs. 32% vs. 31% Baseline Pain (VAS, 0-100): 60.9 ± 20.1 vs. 65.5 ± 18.5 vs. 60.7 ± 21.4 Disability (QDS): 56.0 ± 17.6 vs. 58.6 ± 14.6 vs. 57.4 ± 16.3</p>	<p>A. vs. B. vs. C. Mean improvement in scores from baseline Pain (VAS, 0–100) 1 month (mean): 24.2 (n=80) vs. 25.9 (n=84) vs. 23.4 (n=83) 1 month differences (95% CI) A. vs. B.: -1.7 (NR) A. vs. C: 0.8 (-8.2 to 9.8) 2 months (mean): 37.0 (n=77) vs. 38.1 (n=82) vs. 37.3 (n=79) 2 months difference (95% CI) A. vs. B.: -1.1 (NR) A. vs. C: -0.3 (-9.4 to 10.0) 6 months (mean): 46.8 (n=72) vs. 48.2 (n=78) vs. 47.8 (n=75) 6 months difference (95% CI) A. vs. B.: -1.4 (NR) A. vs. C: -1.0 (-10.0 to 8.0)</p>	<p>A. vs. B. vs. C. Mean improvement in scores from baseline Disability (QDS, 0–100) 1 month (mean): 15.7 (n=80) vs. 11.4 (n=84) vs. 16.2 (n=83) 1 month differences (95% CI) A. vs. B.: 4.3 (NR) A. vs. C: -0.5 (-6.3 to 5.3) 2 months (mean): 26.3 (n=77) vs. 23.5 (n=82) vs. 26.3 (n=79) 2 months difference (95% CI) A. vs. B.: 2.8 (NR) A. vs. C: 0.0 (-7.2 to 7.3) 6 months (mean): 34.6 (n=72) vs. 32.7 (n=78) vs. 35.4 (n=75) 6 months difference (95% CI) A. vs. B.: 1.9 (NR) A. vs. C: -0.7 (-8.4 to 6.9) Cumulative No. of patients, A. vs. B. vs. C; OR (95% CI) Treatment failure 1 month: 2% (n=2) vs. 6% (n=5) vs. 7% (n=6); A. vs. C: 0.3 (0.1–1.6); A. vs. B.: NR 2 months: 13% (n=11) vs. 19% (n=16) vs. 12% (n=10); A. vs. C: 1.1 (0.7–2.8); A. vs. B.: NR 6 months: 23% (n=19) vs. 25% (n=21) vs. 17% (n=14); A. vs. C: 1.5 (0.7–3.2); A. vs. B.: NR Surgery 1 month: 2% (n=2) vs. 5% (n=4) vs. 6% (n=5); A. vs. C: 0.4 (0.1–2.0); A. vs. B: NR 2 months: 12% (n=10) vs. 13% (n=11) vs. 11% (n=9); A. vs. C: 1.1 (0.4–2.9); A. vs. B.: NR 6 months: 16% (n=13) vs. 19% (n=16) vs. 13% (n=11); A. vs. C: 1.2 (0.5–2.9); A. vs. B.: NR</p>

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Hurley, 2015 (26) 52 weeks Chronic Fair	A. Exercise class for 8 weeks (n=83) B. Walking program for 8 weeks (n=82) C. Usual physiotherapy for 8 weeks (n=81)	A. vs. B. vs. C. Mean age: 45.8 vs. 46.2 vs. 44.2 years Female: 71% vs. 71% vs. 62% Mean pain over past week, NRS (0- 10): 5.6 vs. 5.5 vs. 6.0 ODI: 38 vs. 35 vs. 33 EQ-5D: 0.52 vs. 0.57 vs. 0.51	A. vs. B. vs. C. Average pain, NRS (0-10): 5.1 vs. 4.2 vs. 4.1; p=0.15	A. vs. B. vs. C. ODI: 27 vs. 27 vs. 27; p=0.37 EQ-5D: 0.62 vs. 0.63 vs. 0.62; p=0.72
Inani, 2013 (34) 3 months LBP duration not specified Poor	A: MCE; phase 1, patient taught to cognitively perform skilled activation of deep muscle while relaxing superficial muscle; phase 2, improve precision of task including coordinating with breathing, progression to static function position, progression to light dynamic task; phase 3, coordinate the activity of deep and superficial muscles without the global muscle taking over using closed and open chain activities; phase 4 function re-education, subject specific; exercises included transversus abdominus and lumbar multifidus exercises, slow curl-ups, sit-ups, oblique plan/side bridge, and bird-dog exercises.(n=15) B: Conventional exercise; stretching, isometric exercises of spine (hollowing in abdominals, isometric for back extensors), bridging exercises, graded active flexion and extension exercises of spine (n=15) For both groups: 4 weeks regular continuous monitoring in OPD followed by successive followup 3x/wk for remaining 2 months; ergonomic advice given	A. vs. B. Mean age (years): 27.8 vs. 32.9 Female: 40.0% vs. 26.7% Baseline Pain intensity (VAS 0- 10): 6.3 ± 1.8 vs. 7.0 ± 1.6 Function/disability (modified ODI): 19.0 ± 6.4 vs. 21.4 ± 5.4 Disability (%): 38.0 ± 13.0% vs. 42.9 ± 11.0%	A. vs. B. (mean ± SD, t-test) VAS pain (0–10 cm): 1.4 ± 0.9 vs. 2.3 ± 1.1, t=2.273, p=0.031	A. vs. B. (mean ± SD, t-test) Modified ODI: 4.4 ± 2.3 vs. 8.0 ± 3.2, t=3.443, p=0.002 Disability (%): 8.8 ± 4.7% vs. 16.0 ± 6.5%, t=3.443, p=0.002

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Jensen, 2012 (27) 52 weeks Acute, subacute, chronic Good	A. Rest, avoiding hard physical activity and rest twice daily for one hour over 10 weeks (n=50) B. Exercise for 10 weeks (n=50)	A. vs. B. Mean age: 47 vs. 45 years Female: 67% vs. 69% Mean pain, NRS (0- 10): 5.6 vs. 5.1 Mean RDQ: 12.0 vs. 13.3 Mean EQ-5D: 0.68 vs. 0.62 Mean BDI: 10.7 vs. 9.6	A. vs. B. (adjusted differences for intervention group) <u>Post treatment</u> Pain: 5.0 vs. 4.5; adjusted difference -0.07 (95% CI - 0.9 to 0.7) <u>One-year followup</u> Pain: 4.8 vs. 4.3; adjusted difference -0.3 (95% CI - 1.3 to 0.6)	A. vs. B. (adjusted differences for intervention group) <u>Post treatment</u> RDQ: 11.0 vs. 11.1; adjusted difference -0.6 (95% CI -2.2 to 1.0) EQ-5D: 0.7 vs. 0.7; adjusted difference 0.04 (95% CI -0.007 to 0.09) BDI: 8.6 vs. 7.9; adjusted difference 0.67 (95% CI -0.99 to 2.3) vs. 0.08 (95% CI -0.3 to 0.4) <u>One-year followup</u> RDQ: 10.7 vs. 10.7; adjusted difference -1.2 (95% CI -3.3 to 1.0) EQ-5D: 0.7 vs. 0.7; adjusted difference 0.06 (95% CI -0.008 to 0.14) BDI: 9.5 vs. 8.0; adjusted difference -0.92 (95% CI -2.8 to 0.97) vs. -0.17 (95% CI -0.6 to 0.22)

**Author, Year
Duration of
Followup
LBP Duration
Quality**

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Kell, 2011 (28) 13 weeks Subacute, chronic Poor	A. PMR training four days per week with 1,563 repetitions each week (n=60) B. PMR training three days per week with 1,344 repetition each week (n=60) C. PMR training twice per week with 564 repetitions per week (n=60) D. No training (n=60)	A. vs. B. vs. C. vs. D. Mean age: 42.4 ± 5.6 vs. 41.7 ± 6.1 vs. 42.8 ± 6.3 vs. 43.2 ± 5.9 Female: 30% vs. 37% vs. 33% vs. 38.3%	A. vs. B. vs. C. vs. D. VAS pain (0-10): 4.35 ± 0.95 vs. 4.77 ± 1.00 vs. 4.96 ± 1.03 vs. 5.70 ± 0.86 p≤0.05 difference A. vs. B., C, and D p≤0.05 difference B and C vs. D	A. vs. B. vs. C. vs. D. Bench press (function): 79.3 ± 9.7 vs. 70.4 ± 9.1 vs. 68.2 ± 9.7 vs. 53.3 ± 9.3 p≤0.05 difference A. vs. B., C, and D Lat pull down (function): 75.3 ± 7.1 vs. 70.1 ± 7.7 vs. 67.2 ± 7.4 vs. 56.0 ± 6.1 p≤0.05 difference A. vs. B., C, and D p≤0.05 difference B and C Leg press (function): 237.2 ± 29.0 vs. 201.7 ± 30.8 vs. 184.2 ± 29.5 vs. 139.9 ± 28.9 p≤0.05 difference A. vs. B., C, and D p≤0.05 difference B and C ODI: 27.1 ± 10.7 vs. 31.6 ± 11.1 vs. 31.8 ± 10.9 vs. 39.1 ± 10.1 p≤0.05 difference A. vs. B., C, and D p≤0.05 difference B and C vs. D PCS: 55.7 ± 7.8 vs. 50.4 ± 8.0 vs. 50.2 ± 8.7 vs. 45.0 ± 8.0 p≤0.05 difference A. vs. B., C, and D p≤0.05 difference B and C vs. D MCS: 57.7 ± 8.2 vs. 52.6 ± 7.8 vs. 53.1 ± 8.3 vs. 46.0 ± 8.2 p≤0.05 difference A. vs. B., C, and D p≤0.05 difference B and C vs. D

Author, Year
Duration of
Followup
LBP Duration
Quality

	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Little, 2008 (29) 52 weeks Subacute, chronic Good	A. Exercise + 24 lessons in Alexander technique (n=71) B. Exercise + 6 lessons in Alexander technique (n=71) C. Exercise + massage (n=72) D. Exercise (n=72) E. 24 lessons in Alexander technique (n=73) F. 6 lessons in Alexander technique (n=73) G. Massage (n=75) H. Usual care (n=72)	A. vs. B. vs. C. vs. D. vs. E. vs. F. vs. G. vs. H. Mean age: 46 vs. 45 vs. 45 vs. 45 vs. 46 years Female sex: 73% vs. 78% vs. 63% vs. 64% vs. 68% vs. 71%	A. vs. B. vs. C. vs. D. vs. E. vs. F. vs. G. vs. H. Number of days of pain in previous 4 months, difference vs. usual care: -20 (p=0.001) vs. -13 (p=0.031) vs. -11 vs. -11 vs. -20 (p=0.001) vs. -13 (p=0.034) vs. -8 vs. 0 (ref)	A. vs. B. vs. C. vs. D. vs. E. vs. F. vs. G. vs. H. RDQ, difference vs. usual care: -4.22 (p=0.002) vs. -2.98 (p=0.002) vs. -2.37 (p=0.015) vs. -1.65 vs. -4.14 (p<0.001) vs. -1.44 vs. -0.45 vs. 0 (ref) SF-36 PCS, difference vs. usual care: 9.43 (p=0.015) vs. 8.53 (p=0.029) vs. 3.63 vs. -2.08 vs. 11.83 (p=0.002) vs. 2.04 vs. -1.45 vs. 0 (ref) SF-36 MCS, difference vs. usual care: 4.99 vs. 0.64 vs. 2.73 vs. 0.72 vs. 3.74 vs. 4.10 vs. -2.11 vs. 0 (ref)

Author, Year
Duration of
Followup
LBP Duration
Quality

	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Macedo, 2012 (30) 12 months Subacute, chronic Fair	<p>A: MCE; stage 1 = retraining program to improve activity of muscles assessed to have poor control and reduce activity of any muscle identified to be overactive; taught how to contract trunk muscles in a specific manner and progress until able to maintain isolated contractions of the target muscles for 10 reps of 10 seconds each while maintaining normal respiration (feedback available to enhance learning); additional exercises for breathing control, spinal posture, and lower limb and trunk movement were performed; stage 2 = progression toward more functional activities, first using static and then dynamic tasks; motor control exercise guided by pain, and exercises were mostly pain-free. (n=86)</p> <p>B: Graded activity; increase activity tolerance by performing individualized and submaximal exercises (based on activities that each participant identified as problematic/could not perform due to pain), in addition to ignoring illness behaviors and reinforcing wellness behaviors; activities progressed in a time-contingent manner; patients received daily quotas and instructed to only perform the agreed amount. (n=86)</p> <p>Both groups to receive 14 individually supervised sessions of approximately 1 hour (12 initial treatment sessions over an 8-week period [2x wk for first 4 weeks then 1x/wk for next 4 weeks] and 2 booster sessions at 4 and 10 months following randomization; advised to do home exercises (type, intensity, number at discretion of physical therapy) for 30 minutes/wk in first month and 1 hr/wk in second month.</p>	<p>A. vs. B. Mean age: 48.7 vs. 49.6 years Female: 66.3% vs. 52.3% Baseline Pain intensity (NRS 0-10): 6.1 vs. 6.1 Function (PSFS): 3.7 vs. 3.6 Disability (RDQ-24): 11.4 vs. 11.2 Quality of Life (SF-36 PCS and MCS): 43.9 vs. 43.8 and 52.9 vs. 54.7 Global impression of change: -1.4 vs. -1.6</p>	<p>A. vs. B. (mean \pm SD; adjusted treatment effect (95% CI)) Pain intensity (NRS 0-10) baseline: 6.1 \pm 1.9 vs. 6.1 \pm 2.1 (NS) 2 months: 4.1 \pm 2.5 vs. 4.1 \pm 2.5, 0.0 (-0.7 to 0.8), p=0.94 6 months: 4.1 \pm 2.5 vs. 4.1 \pm 2.7, 0.0 (-0.8 to 0.8), p=0.99 12 months: 3.7 \pm 2.7 vs. 3.7 \pm 2.6, 0.1 (-0.7 to 0.9), p=0.83</p>	<p>A. vs. B. (mean \pm SD; adjusted treatment effect (95% CI)) Function (PSFS) baseline: 3.7 \pm 1.6 vs. 3.6 \pm 1.6 (NS) 2 months: 5.9 \pm 2.1 vs. 5.5 \pm 2.4, 0.2 (-0.5 to 0.9), p=0.53 6 months: 5.7 \pm 2.3 vs. 5.7 \pm 2.4, -0.2 (-0.9 to 0.5), p=0.53 12 months: 5.9 \pm 2.2 vs. 6.1 \pm 2.3, -0.4 (-1.1 to 0.3), p=0.25 Disability (RDQ-24) baseline: 11.4 \pm 4.8 vs. 11.2 \pm 5.3 (NS) 2 months: 7.5 \pm 6.4 vs. 8.0 \pm 6.5, -0.8 (-2.2 to 0.7), p=0.30 6 months: 8.0 \pm 7.1 vs. 8.6 \pm 6.8, -0.8 (-2.3 to 0.6), p=0.26 12 months: 7.4 \pm 6.7 vs. 8.0 \pm 6.9, -0.6 (-2.0 to 0.9), p=0.45 Quality of Life, SF-36 PCS baseline: 43.9 \pm 10.8 vs. 43.8 \pm 10.3 (NS) 2 months: 51.6 \pm 12.0 vs. 51.6 \pm 13.4, -0.2 (-13.7 to 3.2), p=0.89 6 months: 52.6 \pm 13.0 vs. 51.2 \pm 13.8, 1.1 (-2.4 to 4.6), p=0.54 12 months: 53.8 \pm 12.7 vs. 53.3 \pm 14.0, -0.3 (-3.8 to 3.3), p=0.88 Quality of Life, SF-36 MCS baseline: 52.9 \pm 10.5 vs. 54.7 \pm 11.5 (NS) 2 months: 56.0 \pm 10.9 vs. 55.8 \pm 13.0, 2.3 (-0.7 to 5.3), p=0.14 6 months: 54.9 \pm 10.4 vs. 56.9 \pm 11.8, 0.1 (-3.0 to 3.1), p=0.97 12 months: 57.0 \pm 10.1 vs. 58.2 \pm 10.8, 0.8 (-2.3 to 3.9), p=0.62 Global impression of change (GPE) baseline: -1.4 \pm 2.3 vs. -1.6 \pm 2.6 (NS) 2 months: 2.0 \pm 1.9 vs. 2.0 \pm 1.9, -0.1 (-1.0 to 0.7), p=0.74 6 months: 1.6 \pm 2.4 vs. 1.5 \pm 2.5, 0.0 (-0.9 to 0.8), p=0.91 12 months: 1.8 \pm 2.5 vs. 1.5 \pm 2.5, 0.2 (-0.6 to 1.0), p=0.62</p>

Author, Year
Duration of
Followup
LBP Duration
Quality

	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Machado, 2010 (31) 3 months Acute <i>Fair</i>	<p>A: McKenzie method + first-line care (n=73). Number of treatment sessions at discretion of the physical therapy, with a max of 6 session over 3 weeks; encouraged to perform the prescribed exercises at home and to follow physical therapists' postural advice at all times; some participants received lumbar support (93%, original McKenzie lumbar roll).</p> <p>B: First-line care only (n=73). Consisted of advice to remain active and to avoid bed rest, reassurance of the favorable prognosis of acute LBP and instructions to take acetaminophen (paracetamol) on a time-contingent basis (NSAIDs not prescribed however those already on them were allow to remain on them); 3 weeks, return for followup as needed during that time</p>	<p>A. vs. B. Mean age: 47.5 vs. 45.9 years Female: 52% vs. 48% Baseline Pain (NRS 0-10): 6.6 ± 1.8 vs. 6.3 ± 1.9 Function (PSFS): 3.7 ± 1.6 vs. 3.4 ± 1.8 Disability (RDQ): 13.7 ± 5.5 vs. 13.5 ± 5.3</p>	<p>A. vs. B. (treatment effects [95% CI] are model-based adjusted differences in outcomes between groups) Pain (NRS 0-10) 1 week: -0.4 (-0.8 to -0.1); p=0.02 (A, n=70; B, n=69) 3 weeks: -0.7 (-1.2 to -0.1); p=0.02 (A, n=70; B, n=68) Mean pain over first 7 days: -0.3 (-0.5 to -0.0); p=0.02 (A, n=70; B, n=69)</p>	<p>A. vs. B. (treatment effects [95% CI] are model-based adjusted differences in outcomes between groups) Function (PSFS) 1 week: 0.0 (-0.4 to 0.5); p=0.90 (A, n=70; B, n=68) 3 weeks: 0.0 (-0.7 to 0.8); p=0.90 (A, n=70; B, n=69) Disability (RDQ) 1 week: -0.2 (-1.5 to 1.0); p=0.74 (A, n=70; B, n=68) 3 weeks: -0.3 (-2.3 to 1.6); p=0.74 (A, n=70; B, n=69) Global perceived effect 1 week: 0.5 (-0.0 to 1.1); p=0.07 (A, n=70; B, n=68) 3 weeks: 0.3 (-0.3 to 0.8); p=0.33 (A, n=70; B, n=69) Development of persistent LBP: 53% (37/70) vs. 47% (32/68); RR 1.1, 95% CI 0.8 to 1.6, p=0.49 Sought additional health care for LBP complaints: 7% (5/70) vs. 26% (18/68); RR 0.27, 95% CI 0.1 to 0.7, p=0.002</p>

**Author, Year
Duration of
Followup
LBP Duration
Quality**

	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Pengel, 2007 (32) 12 months Acute, subacute Fair	<p>A: Exercise and advice (n=63). B: Sham exercise and advice (n=63). C: Exercise and sham advice (n=65). D: Sham exercise and sham advice (n=68).</p> <p>Exercise: Based on program described by Lindstrom and colleagues, to improve the abilities of participants to complete functional activities that they specified as being difficult to perform because of low back pain and includes: aerobic exercise (for example, a walking or cycling program), stretches, functional activities, activities to build speed, endurance, and coordination, and trunk- and limb-strengthening exercises. Physical therapists used principles of cognitive-behavioral therapy and provided individualized home exercise programs.</p> <p>Sham exercise: Sham pulsed ultrasonography (5 minutes) and sham pulsed short-wave diathermy (20 minutes).</p> <p>Advice: Based on the program by Indahl and colleagues and aimed to encourage a graded return to normal activities. Physical therapists explained the benign nature of LBP, addressed any unhelpful beliefs about back pain, and emphasized that being overly careful and avoiding light activity would delay recovery.</p> <p>Sham advice: Participants could talk about their LBP and any other problems, physical therapist responded in a warm and empathic manner, displaying genuine interest, but did not give advice about the LBP.</p> <p>The 12 exercise or sham exercise sessions were delivered over 6 weeks: 3 sessions per week in weeks 1 and 2, 2 sessions per week in weeks 3 and 4, and 1 session per week in weeks 5 and 6. In weeks 1, 2, and 4, participants also received advice or sham advice.</p>	<p>A. vs. B. vs. C. vs. D. D. Mean age (years): 50.1 vs. 51.2 vs. 48.0 vs. 50.0 Female: 46% vs. 44% vs. 46% vs. 54% Baseline Pain (NRS 0-10): 5.4 ± 2.2 vs. 5.5 ± 2.1 vs. 5.4 ± 1.9 vs. 5.3 ± 1.7 Function (PSFS): 3.8 ± 1.9 vs. 3.8 ± 1.8 vs. 3.7 ± 2.0 vs. 4.0 ± 1.7 Disability (RDQ): 9.1 ± 4.8 vs. 8.2 ± 4.4 vs. 8.3 ± 5.0 vs. 8.1 ± 5.6 Global perceived effect: -0.4 ± 2.3 vs. 0.2 ± 2.3 vs. -0.3 ± 2.6 vs. 0.5 ± 2.3 Depression (DASS): 7.3 ± 8.8 vs. 7.4 ± 7.7 vs. 7.1 ± 7.8 vs. 7.1 ± 7.6 Anxiety (DASS): 4.7 ± 6.7 vs. 5.2 ± 7.4) vs. 6.2 ± 7.6 vs. 5.4 ± 6.9 Stress (DASS): 10.1 ± 9.0 vs. 11.7 ± 8.7 vs. 12.6 ± 9.1 vs. 11.7 ± 10.0</p>	<p>A. vs. B. vs. C. vs. D. adjusted multivariable mixed model, relative change (95% CI) <u>Exercise vs. No Exercise</u> Pain (NRS 0-10) 6 weeks: -0.8 (-1.3 to - 0.3), p=0.004 3 months: -0.5 (-1.1 to 0.1), p=0.092 12 months: -0.5 (-1.1 to 0.2), p=0.138 <u>Exercise + Advice vs. No Exercise or Advice</u> Pain (NRS 0-10) 6 weeks: -1.5 (-2.2 to - 0.7), p<0.001 3 months: -1.1 (-2.0 to - 0.3), p=0.009 12 months: -0.8 (-1.7 to 0.1), p=0.069</p>	<p>A. vs. B. vs. C. vs. D. adjusted multivariable mixed model, relative change (95% CI) <u>Exercise vs. No Exercise</u> Function (PSFS) 6 weeks: 0.4 (-0.2 to 1.0), p=0.174 3 months: 0.5 (0.0 to 1.1), p=0.063 12 months: 0.5 (-0.1 to 1.0), p=0.094 Disability (RDQ): 6 weeks: -0.8 (-1.8 to 0.3), p=0.141 3 months: -0.1 (-1.2 to 1.1), p=0.901 12 months: -0.3 (-1.6 to 0.9), p=0.597 Global perceived effect 6 weeks: 0.5 (0.1 to 1.0), p=0.017 3 months: 0.5 (0.1 to 1.0), p=0.030 12 months: 0.4 (-0.1 to 1.0), p=0.134 Depression (DASS) 6 weeks: -0.7 (-2.5 to 1.2), p=0.47 3 months: -0.3 (-2.1 to 1.6), p=0.78 12 months: -0.6 (-2.6 to 1.3), p=0.51 <u>Exercise + Advice vs. No Exercise or Advice</u> Function (PSFS) 6 weeks: 1.1 (0.3 to 1.9), p=0.006 3 months: 1.3 (0.6 to 2.1), p=0.001 12 months: 1.1 (0.3 to 1.8), p=0.005 Disability (RDQ): 6 weeks: -1.3 (-2.7 to 0.2), p=0.085 3 months: -1.0 (-2.6 to 0.6), p=0.20 12 months: -0.9 (-2.7 to 0.8), p=0.29 Global perceived effect 6 weeks: 1.3 (0.7 to 1.9), p<0.001 3 months: 0.8 (0.2 to 1.5), p=0.017 12 months: 0.8 (0.0 to 1.6), p=0.059 Depression (DASS) 6 weeks: 0.2 (-2.5 to 2.8), p=0.91 3 months: 0.2 (-2.4 to 2.7), p=0.91 12 months: -0.4 (-3.1 to 2.3), p=0.76</p>

**Author, Year
Duration of
Followup
LBP Duration
Quality**

Intervention and Duration of Treatment

Population

Pain Outcomes

Other Outcomes

Stankovic, 2012 (33) 4 weeks Chronic <i>Poor</i>	A. Combined exercise program (spinal segmental stabilization plus strengthening and stretching aerobic exercises), 30 minute sessions 5 times a week for 4 weeks (n=100) B. Strengthening and stretching aerobic exercise without pelvic immobilization and core stabilization (n=60)	A. vs. B. Mean age (years): 50 vs. 50 Female: 60% vs. 62% Baseline pain (0-10): 3.1 vs. 3.2 ODS (0-100): 34 vs. 38	A. vs. B. Pain (0-10), mean: 3.13 vs. 3.22 at baseline, 2.03 vs. 2.77 at 4 weeks	A. vs. B. ODS (0-100): 34 vs. 38 at baseline, 23 vs. 33 at 4 weeks
Tai Chi Trials Hall, 2011 (35) 10 weeks Subacute, chronic <i>Fair</i>	A. Tai chi, 18 sessions over 10 weeks (n=80) B. Waitlist (n=80)	A. vs. B. Mean age: 43 vs. 44 years Female: 79% vs. 70%	A. vs. B. Pain, NRS (0-10): 4.4-3.4 vs. 4.4-4.7; mean between-group difference 1.3 (95% CI 0.7 to 1.9) <u>Proportion achieving ≥30% improvement</u> Pain, NRS (0-10): 46.3% vs. 15%; NNT 4	A. vs. B. Bothersomeness, NRS (0-10): 5.0-3.7 vs. 4.5-4.9; mean between-group difference 1.7 (95% CI 0.9 to 2.5) PDI: 22.7-17.0 vs. 23.9-23.8; mean between-group difference 5.7 (95% CI 1.8 to 9.6) RDQ: 10.2-7.0 vs. 9.1-8.1; mean between-group difference 2.6 (95% CI 1.1 to 3.7) QBPDS: 29.2-22.0 vs. 30.2-29.6; mean between-group difference 6.6 (95% CI 2.4 to 10.7) PSFS: 3.5-4.7 vs. 4.0-4.1; mean between-group difference -1.0 (95% CI -1.7 to -0.4) GPE: 0.4-1.6 vs. -0.1-0.4; mean between-group difference -0.8 (95% CI -1.5 to -0.0); p=0.05 <u>Proportion achieving ≥30% improvement</u> Bothersomeness, NRS: 50% vs. 17.5%; NNT 4 PDI, 45% vs. 17.5%; NNT 4 RDQ: 50% vs. 23.8%; NNT 4 QBPDS: 40% vs. 7.5%; NNT 4 PSFS: 43.8% vs. 16.3%; NNT 4
Weifen, 2013 (36) 26 weeks Chronic <i>Fair</i>	A. Tai chi chuan (n=141) B. Backward walking (n=47) C. Jogging (n=47) D. Swimming (n=38)	A. vs. B. vs. C. vs. D. vs. E. Mean age: 37.5 vs. 38.2 vs. 37.2 vs. 37.5 vs. 38.1 years Female: 39% vs.	A. vs. B. vs. C. vs. D. vs. E. VAS (0-100), 3 months: 2.7 vs. 3.3 vs. 3.4 vs. 2.8 vs. 3.6; p<0.05 for A. vs. all other groups except D VAS (0-100), 6 months:	

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
	E. No exercise (n=47)	45% vs. 40% vs. 45% vs. 40% Mean VAS (0- 100): 5.3 vs. 5.2 vs. 5.0 vs. 5.2 vs. 5.1	2.3 vs. 2.9 vs. 3.1 vs. 2.4 vs. 3.2; p<0.05 for A. vs. all other groups except D	
Yoga Trials				
Aboagye, 2015 (44) 1 year Duration not specified	A: Kundalini yoga group class twice a week for 6 week and given CD for home sessions, then encouraged to continue at least twice weekly on own (n=52) B: Exercise therapy with strength training once every wecond week for 6 weeks, then encouraged to continue exercise at least twice weekly on own (n=52) C: Self-care advice: Brief oral recommendatoin to stay active and self-care booklet (n=55)	A vs. B vs. C Mean age: 47 vs. 46 vs. 44 years Female: 72% vs. 62% vs. 80% Baseline pain and function: Not reported	Not reported	A vs. B vs. C Results not reported for entire group Analyzed according to whether intervention performed <2 times/week or ≥2 times/week, adjusted for baseline EQ-5D EQ-5D (mean, 0-1) <2 times/week: 0.72 vs. 0.76 vs. 0.70 at baseline, 0.64 vs. 0.81 vs. 0.74 at 6 weeks, 0.77 vs. 0.81 vs. 0.72 at 6 months, 0.73 vs. 0.75 vs. 0.73 at 12 months (p=0.177 for A vs. B and p=0.073 for A vs. C) ≥2 times/week: 0.71 vs. 0.74 vs. 0.73 at baseline, 0.80 vs. 0.78 vs. 0.70 at 6 weeks, 0.77 vs. 0.76 vs. 0.70 at 6 months, 0.79 vs. 0.79 vs. 0.75 at 12 months (p=0.57 for A vs. B and p=0.031 for A vs. C)
Morone, 2015 (46) 3-4 months Acute to chronic	A: Yoga therapy with postural, breathing, and relaxation exercises 2 or more classes per week for 2 weeks, then 15-30 home session daily (n=30) B: No yoga (n=31)	A vs. B Mean age: 36 vs. 37 years Female: 50% vs. 45% Worst pain in past 2 weeks severe: 23% vs. 19% Baseline RDQ: 12 VS. 11	Not reported	A vs. B RDQ: 9.0 vs. 11.3 at 3-4 m, unadjusted difference 2.5 (95% CI -0.10 to 4.8), adjusted for baseline RDQ, sex, and age 3.29 (95% CI 0.98 to 5.6) Aberdeen Low Back Pain Score: 16.4 vs. 17.0 at 3-4 m, unadjusted difference 0.68 (95% CI - 3.5 to 4.9), adjusted for baseline RDQ, sex and age 1.63 (95% CI -2.4 to 5.6)

Author, Year
Duration of
Followup
LBP Duration
Quality

Intervention and Duration of Treatment

Population

Pain Outcomes

Other Outcomes

Nambi, 2014 (38) 6 months Subacute, chronic <i>Fair</i>	<p>A: Iyengar yoga 1 hour class/week + 30 minute home practice, 5 days/week for 4 weeks; with props; 29 poses introduced in stages simple to progressively more challenging; At end of 4 weeks, participants encouraged to continue Yoga at home (n=30)</p> <p>B: Following 5-10 minute warm up (stretching exercises for soft tissue flexibility and range of motion); Taught specific exercises for strengthening abdominal and back muscles (depending on clinical findings) 3 days/week with 5 repetitions in 3 sets with 30-s pause per set; repetitions gradually increased until reaching 15 for 4 weeks: instructed to refrain from other back exercises, strenuous activities outside of normal activities of daily living during study (n=30)</p>	<p>A. vs. B. Mean age: 44.26 vs. 43.66 Female: 63.34% vs. 43.34% Baseline Pain intensity (10 cm VAS, 0=no pain, 10=worst possible): 6.7 vs. 6.7 Physically unhealthy days (from CDC HRQOL-4): 18 vs. 17.8 Mentally unhealthy days (from CDC HRQOL-4): 17.0 vs. 17.4 Activity limitation days (from CDC HRQOL-4): 16.7 vs. 17.1</p>	<p>A. vs. B. Pain intensity (10 cm VAS, mean): 4 weeks 3.8 vs. 5.3; 6 months 1.8 vs. 3.8, % improvement 72.81% vs. 42.5%, p=0.001; SMD* 4 weeks (-1.66, 95% CI -2.24 to -1.07); 6 months (-2.17, 95% CI -2.81 to -1.53)</p>	<p>A. vs. B. Physically unhealthy days (mean): 4 weeks 7.7 vs. 12.0; 6 months 2.6 vs. 6.9, % improvement 85.61% vs. 61.0%, p=0.001; Mentally unhealthy days (mean): 4 weeks 8.4 vs. 10.5; 6 months 2.6 vs. 6.9, % improvement 87.53% vs. 71.37%, p=0.001; Activity limitation days (mean): 4 weeks 7.5 vs. 12.0; 6 months 2.0 vs. 5.0, % improvement 87.83% vs. 70.59%, p=0.001;</p>
------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Author, Year
Duration of
Followup
LBP Duration
Quality

	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Saper, 2013 (39) Chronic Fair	A: 75 minute Hatha Yoga class once per week + recommended 30 minute home practice (n=49) B: 75 minute Hatha Yoga class twice per week + recommended 30 minute home practice (n=46) 12 weeks	A. vs. B. Mean age: 46.4 vs. 48.7 years Female: 71% vs. 80% Baseline pain (mean, low back pain intensity, 11 point numeric scale) 7.1 vs. 6.7 Back-specific function: (mean RDQ) 13.7 vs. 13.6 SF-26 Physical: 37.5 vs. 37.4; Mental 44.8 vs.44.	A. vs. B. Change from baseline, <i>between group</i> difference in means: Pain: 6 weeks, -0.3 (-1.1 to 0.6), p=0.49; 12 weeks, 0.3 (-0.2 to 0.8), p=0.62 Pain: proportion experiencing ≥30% improvement from baseline: 29% (23/47) vs. 59%(26/44), p=0.33, RR 0.83 (95% CI 0.57 to 1.12); proportion experiencing ≥50% improvement from baseline: 57% (27/47) vs. 66% (29/44), p=0.41, RR 1.14 (95% CI 0.64 to 2.02;	A. vs. B. Change from baseline, <i>between group</i> difference in means: RDQ: 6 weeks -0.6 (-2.7 to 1.6), p=0.62; 12 weeks, -0.1 (-1.4 to 1.2), p=0.83 RDQ proportion experiencing ≥30% improvement from baseline: 57% (27/47) vs. 66%(29/44), p=0.41, RR 0.87 (95% CI 0.63 to 1.21): proportion experiencing ≥50% improvement from baseline: 47% (22/47) vs. 50% (22/44), p=0.76, RR 0.94 (95% CI 0.61 to 1.43) Change from baseline, between group difference in means SF-36 Physical: 6 weeks 1.6 (95% CI -1.6 to 4.9) p=0.33; 12 weeks 0.2 (-3.4 to 3.7) p=0.93; SF-36 Mental 6 weeks 2.2 (-1.9 to 6.3) p=0.29; 12 weeks 1.5 (-2.6 to 5.6) p=0.47. Overall improvement scores: Same for A and B (mean 4.5, median 5) Satisfaction scores: mean 1.3 vs. 1.5, median 1 for both Medication use: Use of any pain medication decrease at 6 weeks (27% vs. 35%) and remained similar at 12 weeks, but NS difference in use of any pain medication or specific analgesic categories. Per protocol analyses did not reveal any statistical differences between groups for any outcome; Dose-response: Substantial variability in data; authors report potential for a "modest" dose-response" relationship with decrease in relationship slope for change in pain at approximately 12 class and approximately 9 classes for RDQ -figure provided, but not detailed data -Authors indicated that conclusions regarding the causality of the association are not possible. Adherence: Class attendance: 65% (32/47) vs. 44% (20/44), p=0.04; weekly amount of home practice 93 vs. 97 minutes; home practice for both groups a median of 4 days/week; Hours of class + home 37 vs. 29, p=0.037

Author, Year
Duration of
Followup
LBP Duration
Quality

Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Mindfulness-based Stress Reduction Trials				
Cherkin, 2016 (45) 52 weeks Chronic Good	A: Mindfulness-based stress reduction via group sessions (2 hours per week for 8 weeks, with optional 6 hour retreat) (n=116) B: Cognitive behavioral therapy via group sessions (2 hours per week for 8 weeks) (n=113) B: Usual care (n=113)	A. vs. B. vs. C. Mean age: 50 vs. 49 vs. 49 years Female: 61% vs. 59% vs. 77% Pain bothersomeness: 6.0 vs. 5.8 vs. 5.8 RDQ: 11.8 vs. 11.5 vs. 0.9	A. vs. B. vs. C. ≥30% improvement in pain bothersomeness: 19% vs. 22% vs. 21% at 4 weeks, RR 0.93 (95% CI 0.56 to 1.52) for A vs. C; 36% vs. 34% vs. 25% at 8 weeks, RR 1.46 (95% CI 0.99 to 2.16) for A vs. C; 44% vs. 45% vs. 27% at 26 weeks, RR 1.64 (95% CI 1.15 to 2.34) for A vs. C; 48% vs. 40% vs. 31% at 52 weeks, RR 1.56 (95% CI 1.14 to 2.14) for A vs. C	A. vs. B. vs. C. ≥30% improvement in RDQ: 34% vs. 25% vs. 27% at 4 weeks, RR 1.26 (95% CI 0.86 to 1.86) for A vs. C; 47% vs. 52% vs. 35% at 8 weeks, RR 1.34 (95% CI 0.98 to 1.84) for A vs. C; 60% vs. 58% vs. 44% at 26 weeks, RR 1.37 (95% CI 1.06 to 1.77) for A vs. C; 69% vs. 59% vs. 49% at 52 weeks, RR 1.41 (95% CI 1.13 to 1.77) for A vs. C RDQ (mean change from baseline, 0 to 24): -1.93 vs. -1.44 vs. -1.28 at 4 weeks, difference -0.65 (95% CI -1.59 to 0.28) for A vs. C; -3.40 vs. -3.37 vs. -1.83 at 8 weeks, difference -1.57 (95% CI -2.70 to -0.45) for A vs. C; -4.33 vs. -4.38 vs. -2.96 at 26 weeks, difference -1.37 (95% CI -2.55 to -0.19) for A vs. C; -5.3 vs. -4.78 vs. -3.43 at 52 weeks, difference -1.87 (95% CI -3.14 to -0.60) for A vs. C 1.26 (95% CI -0.60 to 3.11) Global improvement (pain much better or completely gone): 16% vs. 22% vs. 11% at 8 weeks, RR 1.45 (95% CI 0.76 to 2.78) for A vs. C; 26% vs. 30% vs. 14% at 26 weeks, RR 1.93 (95% CI 1.12 to 3.32) for A vs. C; 30% vs. 32% vs. 18% at 52 weeks, RR 1.67 (95% CI 1.03 to 2.71) for A vs. C Used medication in past week: 53% vs. 53% vs. 63% at 8 weeks, RR 0.84 (95% CI 0.70 to 1.02) for A vs. C; 43% vs. 51% vs. 54% at 26 weeks, RR 0.80 (95% CI 0.63 to 1.02) for A vs. C; 47% vs. 42% vs. 53% at 52 weeks, RR 0.89 (95% CI 0.70 to 1.11) for A vs. C

Author, Year
Duration of
Followup
LBP Duration
Quality

Intervention and Duration of Treatment

A: Mindfulness-based stress reduction via group sessions (2 hours per week for 8 weeks, with optional 6 hour retreat) (n=116)

B: Cognitive behavioral therapy via group sessions (2 hours per week for 8 weeks) (n=113)

B: Usual care (n=113)

Population

Pain Outcomes

Other Outcomes

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Cherkin, 2016 (Continued) 52 weeks Chronic Good	A: Mindfulness-based stress reduction via group sessions (2 hours per week for 8 weeks, with optional 6 hour retreat) (n=116) B: Cognitive behavioral therapy via group sessions (2 hours per week for 8 weeks) (n=113) B: Usual care (n=113)		Pain bothersomeness (mean change from baseline, 0 to 10 scale): -0.57 vs. -0.79 vs. -0.68 at 4 weeks, difference -0.65 (95% CI -1.59 to 0.28) for A vs. C; -1.40 vs. -1.28 vs. -0.67 at 8 weeks, difference -0.73 (95% CI -1.19 to -0.27) for A vs. C; -1.48 vs. -1.56 vs. -0.84 at 26 weeks, difference -0.64 (95% CI -1.18 to -0.11) for A vs. C; -1.95 vs. -1.76 vs. -1.10 at 52 weeks, difference -0.85 (95% CI -1.39 to -0.32) for A vs. C	PHQ-8 (mean change from baseline, 0 to 24 scale): -1.60 vs. -2.29 vs. -0.12 at 8 weeks, difference -1.48 (95% CI -2.31 to -0.64) for A vs. C; -1.32 vs. -1.80 vs. -0.64 at 26 weeks, difference -0.68 (95% CI -1.45 to 0.09) for A vs. C; -1.51 vs. -1.72 vs. -0.88 at 52 weeks, difference -0.62 (95% CI -1.48 to 0.23) GAD-2 (mean change from baseline, 0 to 6 scale): -0.33 vs. -0.51 vs. -0.09 at 8 weeks, difference -0.24 (95% CI -0.56 to 0.09); 0.00 vs. -0.49 vs. 0.02 at 26 weeks, difference -0.02 (95% CI -0.41 to 0.37); -0.15 vs. -0.39 vs. -0.14 at 52 weeks, difference 0.00 (95% CI -0.37 to 0.36) SF-12, Physical Component Score (mean change from baseline, 0 to 100 scale): 3.69 vs. 3.24 vs. 2.21 at 8 weeks, difference 1.48 (95% CI -0.06 to 3.02); 3.58 vs. 3.78 vs. 3.27 at 26 weeks, difference 0.31 (95% CI -1.53 to 2.16); 3.87 vs. 3.79 vs. 2.93 at 52 weeks, difference 0.94 (95% CI -0.86 to 2.74) SF-12, Mental Component Score (mean change from baseline, 0 to 100 scale): 1.68 vs. 1.77 vs. -0.65 at 8 weeks, difference 2.33 (95% CI 0.68 to 3.99); 0.45 vs. 2.13 vs. -1.11 at 26 weeks, difference 1.57 (95% CI -0.27 to 3.40); 2.01 vs. 1.81 vs. 0.75 at 52 weeks, difference

Author, Year
Duration of
Followup
LBP Duration
Quality

Intervention and Duration of Treatment

Population

Pain Outcomes

Other Outcomes

Morone, 2016 (46) 6 months Chronic Good	A: Mindfulness-based stress reduction via 8 90 minute weekly group sessions, with booster session every 6 months (n=140) B: Health education with 8 weekly sessions, based on the "10 Keys" to Health Aging (n=142)	A. vs. B. Mean age: 75 vs. 74 years Female: 66% vs. 66% Pain intensity (0-20 NRS): 11.0 vs. 10.5 RDQ (mean, 0-24): 15.6 vs. 15.4	A. vs. B. Average pain (mean, 0 to 20 NRS): 11.0 vs. 10.5 at baseline, 9.6 vs. 9.7 at 8 w, 9.5 vs. 10.6 at 6 m (difference -1.1, 95% CI -2.2 to -0.01) ≥30% improvement in average pain: 38% vs. 29% at 8 w (p=0.12), 37% vs. 27% at 6 m (p=0.09)	A. vs. B. RDQ (mean, 0-24): 15.6 vs. 15.4 at baseline, 12.1 vs. 13.1 at 8 months (difference -1.1, 95% CI -2.1 to -0.01), 12.2 vs. 12.6 at 6 months (NS) ≥2.5 point improvement in RDQ: 57% vs. 45% at 8 w (p=0.051), 49% vs. 49% at 6 months (p=0.97) SF-36 Global Health Composite (mean, 9 to 67 scale): 40.5 vs. 40.6 at baseline, 42.9 vs. 41.3 at 8 weeks, 42.4 vs. 42.1 at 6 months (NS) SF-36 Physical Health Composite (mean, 20 to 65 scale): 38.8 vs. 38.9 at baseline, 42.1 vs. 40.7 at 8 weeks, 42.4 vs. 42.1 at 6 months (NS) Global impression of change "much improved": 31% (5/16) vs. 11% (2/18)
Morone, 2009 (47) 4 months Chronic Poor	A: Mindfulness-based stress reduction via 8 90- minute weekly group sessions (n=16) B: Health education with 8 90- minute weekly sessions (n=19)	A. vs. B. Mean age: 78 vs. 73 years Female: 69% vs. 58% Average pain (mean, 0-100 SF-36 Pain Score): 40 vs. 40 RDQ (mean, 0-24): 9 vs. 11	A. vs. B. SF-36 Pain Score (mean, 0 to 100): 40 vs. 40 at baseline, 42 vs. 40 at 8 weeks, 41 vs. 40 at 4 months McGill Pain Questionnaire, total score (mean, 0 to 78): 16 vs. 16 at baseline, 12 vs. 11 at 8 weeks, 12 vs. 12 at 4 months McGill Present Pain Intensity (mean, 0 to 5): 2.9 vs. 4.3 at baseline, 2.7 vs. 4.0 at 8 weeks, 2.3 vs. 3.8 at 4 months	A. vs. B. RDQ (mean, 0 to 24): 9 vs. 11 at baseline, 7.5 vs. 9 at 8 weeks, 7.5 vs. 10 at 4 months SF-36 Role Limitations (mean, 0 to 100): 33 vs. 30 at baseline, 35 vs. 26 at 8 weeks, 35 vs. 29 at 4 months
Psychological Therapies Khan, 2014 (49) Post-treatment Sub-acute, chronic	A: Behavioral therapy plus exercise (n=27). Physical-therapist guided sessions 3 times per week for 12 weeks; patients instructed to continue exercises at home twice a day at least 5 times a week. Cognitive behavioral	A. vs. B. Mean age: 40 years Female: 54% Baseline pain (0-	A. vs. B. Pain (mean 0-10 VAS): 6.5 vs. 7.0 at baseline; 2.7 vs. 5.3 post-treatment (p<0.0001)	A. vs. B. Function (mean 0-24 RDQ): 13.8 vs. 12.9 at baseline; 5.3 vs. 9.9 post-treatment (p<0.0001)

**Author, Year
Duration of
Followup
LBP Duration
Quality**

Intervention and Duration of Treatment

Population

Pain Outcomes

Other Outcomes

Fair	therapy aimed to guide patients to achieve their daily life goals, consisting of operant behavioral graded activity and problem solving training. B: Exercise (n=27). Physical-therapist guided sessions 3 times per week for 12 weeks; patients instructed to continue exercises at home twice a day at least 5 times a week. Graded activity led by physical therapist that focused on gradual increase or pacing of activities important for individual patients with general exercises.	10 VAS): 6.5 vs. 7.0 (mean) (p=0.1877) Baseline function (0-24 RDQ):13.8 vs. 12.9 (mean) (p=0.1842)		
Lamb, 2010 (50), 2012 (51) 3, 6, 12 months, and at a mean of 34 months Subacute to chronic Fair	A. Group cognitive behavioral therapy plus active management advisory consult (n=468) B. Active management advisory consult alone (patients free to seek additional care) (n=233) Treatment protocols: CBT (group A): 7x90 minute sessions; treatment duration not reported Active management advisory consult (both groups): 1x15 minute session	A. vs. B. Mean age 53 vs. 54 years 59% vs. 61% female Pain (Van Korf pain): 59 vs. 59 Function (RDQ): 9 vs. 9 Function (Von Korff disability): 49 vs. 46 Quality of life (EQ-5D): not reported Quality of life (SF-12 physical): 37 vs. 38 Quality of life (SF-12 mental): 45 vs. 46 Pain Self-efficacy: 40 vs. 41 Fear avoidance beliefs (Fear	A. vs. B. Pain (mean change from baseline, 0-100% Von Korff pain): 12.2 vs. 5.4 at 3 months (p<0.0001), 13.7 vs. 5.7 at 6 months (p<0.0001), 13.4 v. 6.4 at 12 months (p<0.0001), 17.4 vs. 12.8 at 34 months (p=0.107) Pain self-efficacy (mean change from baseline 0-60 Pain Self Efficacy): -2.4 vs. 0.9 at 3 months (p<0.0001), -2.6 vs. 1.5 at 6 months (p<0.0001), -3.0 vs. 0.8 at 12 months (p<0.0001)	A. vs. B. Function (mean change from baseline, 0-24 RDQ): 2.0 vs. 1.1 at 3 months (p=0.0021), 2.5 vs. 1.0 at 6 months (p=0.0002), 2.4 vs. 1.1 at 12 months (p=0.0008), 2.9 vs. 1.6 at 34 months (p=0.013) Function (mean change from baseline, 0-100% Von Korff disability): 13.2 vs. 8.9 at 3 months (p=0.0316), 13.9 vs. 5.7 at 6 months (p<0.0001), 13.8 vs. 5.4 at 12 months (p<0.0001), 16.7 vs. 11.2 at 34 months (p=0.039) Quality of life (mean change from baseline, -0.59 to 1 EQ-5D): -0.06 vs. 0.01 at 3 months (p=0.007), -0.05 vs. -0.03 at 6 months (p=0.382), -0.06 vs. -0.0003 at 12 months (p=0.027), -0.07 vs. -0.04 at 34 months (p=0.387) Quality of life (mean change from baseline, 0-100 SF-12 physical): -3.7 vs. -1.5 at 3 months (p=0.0031), -3.6 vs. -1.8 at 6 months (p=0.0144), -4.9 vs. -0.8 at 12 months (p<0.0001) Quality of life (mean change from baseline 0-100 SF-12 mental): -1.3 vs. 0 at 3 months (p=0.1276), -2.5 vs. 0.09 at 6 months (p=0.0035), -0.9 vs. -0.7 at 12

Author, Year
Duration of
Followup
LBP Duration
Quality

Intervention and Duration of Treatment

Population

Pain Outcomes

Other Outcomes

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
		avoidance beliefs questionnaire): 14 vs. 14		months (p=0.8323) Treatment benefit (% patients who considered themselves recovered): 59% (235/395) vs. 31% (62/197) at 12 months (p<0.0001) Treatment satisfaction (% patients satisfied with treatment): 65% (212/328) vs. 28% (43/151) at 12 months (p=0.463) Fear avoidance beliefs (mean change from baseline 0-24 Fear Avoidance Beliefs Questionnaire): 3.4 vs. 0.7 at 3 months (p=0.0004), 3.0 vs. -0.1 at 6 months (p<0.0001), 3.4 vs. 0.5 at 12 months (p<0.0001)
Siemonsma, 2013 (52) Post-treatment Chronic Fair	A. Cognitive treatment of illness perceptions (n=104) B. Wait list control (no interventions, could be treated as group A at end of 18 weeks) (n=52) Treatment protocol (group A): 10-14x60 minute sessions over 18 weeks	A. vs. B. Mean age 45 vs. 47 years 51% vs. 60% female Activity-specific pain (PSC): ~76 vs. ~70 (estimated from graph) Function (QBPDS): 40.4 vs. 40.3	A. vs. B. Activity-specific pain (mean 0 to 100 PSC): ~44 vs. ~64 post-treatment (estimated from graph) Activity-specific pain (mean change from baseline, 0 to 100 PSC): -19.1 (95% CI -24.3 to -13.9) vs. -5.2 (95% CI -14.7 to 4.2) (p=0.018) post-treatment Activity-specific pain (% of patients with clinically relevant change, defined as decrease of 18 to 24 mm): 49% (46/93) vs. 26% (12/46) post-treatment (OR 2.77 (95% CI 1.28 to 6.01))	A. vs. B. Function (mean 0-100 QBPDS): 36.9 vs. 38.7 post-treatment (p=0.27)
Vong, 2011 (53) 1 month Chronic Fair	A. Motivational enhancement treatment during physical therapy (n=45) B. Physical therapy (n=20) Treatment protocol: 10x30 minute sessions over 8 weeks	A. vs. B. Mean age 45 vs. 45 years 58% vs. 68% female Pain (0-10 VAS):	A. vs. B. Pain (mean 0-10 VAS): 3.1 vs. 3.9 at 1 month (p>0.05) Pain self-efficacy (mean 0-60 PSEQ): 45.4 vs. 45.6 at 1 month (p>0.05)	A. vs. B. Function (mean 0-24 RDQ): 5.6 vs. 7.6 at 1 month (p>0.05) Function (mean 0-100 SF-36 physical function): p> 0.05 at 1 month (data not reported)

Author, Year
Duration of
Followup
LBP Duration
Quality

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
		5.3 vs. 5.3 Pain self-efficacy (PSEQ): 39.5 vs. 40.5 at baseline Pain (SF-36 bodily pain): 41 vs. 49 (p=0.047) Function (RDQ) (mean): 10.0 vs. 10.0 Function (mean 0-100 SF-36 physical function): 67 vs. 63 Quality of life (SF-36 role-physical): 22 vs. 30 Quality of life (SF-36 general health): 41 vs. 49	Pain (mean 0-100 SF-36 bodily pain): p> 0.05 at 1 month (data not reported)	Quality of life (mean 0-100 SF-36 role-physical and general health scales): p> 0.05 at 1 month (data not reported)
Multidisciplinary Rehabilitation Trials Eisenberg, 2012 (55) 2, 5, 12, and 26 weeks LBP duration not specified Good	A Integrative Care (acupuncture, chiropractic, internal med consult, massage, occupational therapy, physical therapy, mind-body techniques, neuro consult, nutrition counseling, ortho consult, psych and rheum consult as needed) + usual care B. Usual care (medical care) 12 weeks	A. vs. B. Mean age: 47 vs. 48 years Female: 50% vs. 67% Average Pain (0-10): 4.8 vs. 5.7 Modified RDQ: 15.7 vs. 16	A. vs. B. Pain (0-10 scale) Week 2: 3.6 vs. 4.8 (p=0.62) Week 5: 1.9 vs. 5.5 (p=0.05) Week 12: 0.6 vs. 5.0 (p=0.005) Week 26: 1.0 vs. 4.7 (p=0.04)	A. vs. B. RDQ mean differences, A. vs. B. Week 2: 12 vs. 11.3 (p=0.87) Week 5: 8.5 vs. 13 (p=0.26) Week 12: 3.9 vs. 11 (p=0.08) Week 26: 4.3 vs. 10.7 (p=0.10) SF-12 physical Week 2: 35 vs. 41 (p=0.90) Week 5: 42 vs. 42 (p=0.38) Week 12: 49 vs. 43 (p=0.06) Week 26: 51 vs. 44 (p=0.03) SF-12 mental Week 2: 47 vs. 51 (p=0.26) Week 5: 51 vs. 50 (p=0.59) Week 12: 50 vs. 51 (p=0.48)

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
				Week 26: 54 vs. 51 (p=1.00)
				Days in bed, days at home and reduced activity days NS
				Regression showed positive differences significant for RDQ, pain, and bothersomeness at 12 weeks, but not at 26 weeks
Gatchel, 2003 (56) 3,6,9,12 months Acute Fair	A. Intensive Multidisciplinary rehabilitation (physician evaluation, psychology, physical therapy, biofeedback, case management, occupational therapy) B. Usual care	Mean age: 38 years Female: 35% Baseline pain: not reported Baseline function: not reported	A. vs. B. Average self-rated pain over last 3 months: 27 vs. 43, p=0.001	A. vs. B. Return to work at 12 months: 91% vs. 69%, OR 4.55 (p=0.027) Average number of disability days due to back pain: 38 vs. 102, p=0.001 Taking opioid analgesics: 27% vs. 44%, OR 0.44, p=0.020 Cost: \$12,721 vs. \$21,843, p<0.05
Monticone, 2014 (57) 0, 8 weeks; 3 months Chronic Good	A. Multidisciplinary rehabilitation of 2 months duration (physiatry, psychology, occupational therapy, and physiotherapy) providing spinal stabilization and cognitive behavioral therapy to address fear avoidance B. Usual care = passive spinal mobilization, stretching, muscle strengthening, and posture control	A. vs. B. Mean age: 59 vs. 57 years Female: 70% vs. 40% Baseline pain: not reported BMI: 27 vs. 25		A. vs. B. Baseline ODI 26 vs. 24 (p=0.43) TSK 29 vs. 27 (p=0.55) NRS (0-10) 5 vs. 4 (p=0.67) PCS 25 vs. 23 (0.43) SF-36 Physical Activity 41 vs. 43 (p=0.55) 6 minute walk test 1.17 m/s vs. 1.26 m/s (p=0.29)
				8 weeks ODI 10 vs. 8 (p=0.03) TSK 29 vs. 27 (p=0.01) NRS (0-10) 5 vs. 4 (p=1.0) PCS 25 vs. 23 (p=0.006) SF-36 Physical Activity 41 vs. 43 (p=0.001) 6 minute walk test 1.17 m/s vs. 1.26 m/s (p=0.478)
				3 months ODI 8 vs. 15 TSK 15 vs. 27 NRS (0-10) 2 vs. 3 PCS 9 vs. 18

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Acupuncture Trials				
Cho, 2013 (60) Primary: 8 weeks FU to 6 months Chronic Good	A: Acupuncture (n=65) B: Sham acupuncture (n=65) Treatment protocol: 2x weekly x 6 weeks	A. vs. B. Mean age 42 vs. 42 83% vs. 86% female Race not reported Pain, VAS (0-10) 6.5 vs. 6.4 Disability, ODI: 28.2 vs. 24.2	A. vs. B. Pain, bothersomeness (primary) mean change from baseline (0-10 VAS): -3.4 vs. -2.3 (p<0.05) Pain intensity mean change from baseline (0- 10 VAS): -3.5 vs. -2.3 (p=0.008)	SF-36 Physical Activity 84 vs. 67 6 minute walk test 1.53 vs. 1.42 A. vs. B.(to primary endpoint) Disability, Proportion of ODI improvement from baseline: -0.42 vs. 0.29 (NS)
Hasegawa, 2014 (61) 28 days Acute Good	A. Scalp acupuncture +diclofenac (n=40) B. Sham scalp acupuncture +diclofenac (n=40) Treatment protocol: 5 30 min sessions (unclear time period)	A. vs. B. Mean age 47 vs. 44 years 63% vs. 65% female 63% vs. 55% Caucasian Pain, VAS (0-10): 6.6 vs. 6.7 Disability, RDQ: 14.9 vs. 14.6	A. vs. B.: Acute LBP Pain, VAS(0-10) mean change from baseline: - 4.6 vs. -3.3; p=0.005	A. vs. B. Disability, RDQ mean change from baseline: - 10.8 vs. -8.6; p=0.002
Vas, 2012 (62) Primary: 3 weeks FU to 48 weeks Acute Good	A. True acupuncture (n=68) B. Sham acupuncture (n=68) C. Placebo acupuncture (n=69) D. Control group (n=70) Treatment protocol: 5 20 min sessions over 2 weeks	A. vs. B. vs. C. vs. D Mean age 42 vs. 44 vs. 44 vs. 41 63% vs. 57% vs. 49% vs. 64% female Race not reported (Spain)	A. vs. B. vs. C. vs. D Pain VAS not reported Continuing pain and recurrence of pain reported only	A. vs. B. vs. C. vs. D Disability (Proportion achieving 35% improvement in RDQ (0-24) at 3 weeks): 74% vs. 75% vs. 65% vs. 44% (p<0.05 for A. vs. C. and A. vs. D)
Weiss, 2013 (63) 3 months after end of	A. Acupuncture plus intensive rehab (n=74) B. Intensive inpatient rehab only (n=69) Treatment protocol: Daily acupuncture for 21	A. vs. B. Mean age 49.8 vs. 51.7 27% vs. 39.1%	A. vs. B. Bodily pain, SF-36 mean change from baseline to 3 months post treatment 8.3	A. vs. B. Physical function, SF-36 mean change from baseline to 3 months post treatment -3.6 vs. - 11.8 p=0.0.02

**Author, Year
Duration of
Followup
LBP Duration
Quality**

Intervention and Duration of Treatment

Population

Pain Outcomes

Other Outcomes

treatment Chronic Poor	days of inpatient rehab	female Race not reported (Germany) Bodily Pain, SF-36 41.2 vs. 36.0 Physical function, SF-36 71.2 vs. 69.8	vs. 3.8 p=0.28 (p<0.05) Bodily pain, SF-36 mean change from baseline to end of treatment 24.5 vs. 22.6 p=0.56	Physical function, SF-36 mean change from baseline to end of treatment 9.8 vs. 6.4 p=0.20
Yeh, 2016 (69) Chronic 1 month after end of treatment Poor	A. Auricular acupuncture (n=31) B. Sham: Vaccaria seeds taped onto the stomach, mouth, duodenum, and eye acupoints of both ears, patients instructed to press the seeds on each ear (n=30) Treatment protocol: Acupuncture once weekly for 4 weeks; sham at least 3 times a day for 3 minutes and when having pain for 5 days then off for 2 days, total 4 weeks	A. vs. B. Mean age 61 vs. 66 67% vs. 68% female 87% vs. 81% White 13% vs. 19% Black Worst pain: 6.9 vs. 7.0	A. vs. B. Worst pain (0-10 NRS), mean change from baseline: -3.53 vs. -0.77 at 4 weeks, difference - 2.76 (95% CI -3.93 to - 1.59) and -3.70 vs. -2.86 at 8 weeks, difference - 2.86 (95% CI -4.03 to - 1.69)	A. vs. B. Global Pittsburgh Sleep Quality Index (0 to 21), mean change from baseline: -1.42 vs. -0.37 at 4 weeks, difference -1.05 (95% CI -2.22 to 0.13) and -1.22 vs. 0.21 at 8 weeks, difference -1.43 (95% CI -2.60 to -0.25)
Yun, 2012 (64) Chronic 24 weeks Fair	A. Back-pain-acupuncture (n=80) B. Standard acupuncture (n=82) C. Usual care (n=74) Treatment protocol: 14 daily treatments	A. vs. B. vs. C Mean age 33 vs. 34 vs. 31 33% vs. 27% vs. 31%female Race not reported (China) Pain, (0-10 VAS): 6.1 vs. 6.1 vs. 6.1 Disability, RDQ: 11.8 vs. 12 vs. 11.8	A. vs. B. vs. C Pain, bothersomeness (primary) mean change from baseline 24 weeks (0-10 VAS): 2.5 vs. 2.0 vs. 1.2 (p<0.0001)	A. vs. B. vs. C RDQ mean change from baseline: 6.2 vs. 5.3 vs. 4.1 (p<0.0001)
Massage Trials Ajimsha, 2014 (76) 12 weeks Subacute, chronic Fair	A. Myofascial release+ specific back exercise (n=38) B. Sham myofascial release + specific back exercise (n=36) Treatment given 3 times weekly for 8 weeks	A. vs. B. Mean age: 35.8 vs. 34.2 Female: 76% vs. 78% Baseline pain: not reported Baseline function:	A. vs. B. Mean differences, B vs. A: MPQ, week 8: 4.813, p=0.000 MPQ, week 12: 3.25, p=0.000	A. vs. B. Mean differences, B vs. A: QBPDS, 8 weeks: 3.413, p=0.000 QBPDS, 12 weeks: 2.023, p=0.000

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
		not reported		
Borges, 2014 (77) 6 weeks LBP duration: not reported Fair	A. Massage by accupressure (n=14) B. Laser applied but turned off (placebo) (n=15) C. No treatment (n=14)	A. vs. B. vs. C. Mean age: 39.6 overall Female: 92.9% vs. 73.3% vs. 64.3 Pain score: of 7: 64.3% vs. 26.7% vs. 21.4% Baseline function: not reported	A. vs. B. vs. C Pain scores, baseline vs. 3 weeks vs. 6 weeks: A: 6.4 vs. 3.4 vs. 0.9, p<0.001 B: 5.7 vs. 4.8 vs. 4.7, p>0.05 C: 5.0 vs. 5.3 vs. 5.9, p>0.05	
Cherkin, 2011 (71) 1 month and 3 months >12 weeks Chronic Good	A. Structural massage (n=132) B. Relaxation massage (n=136) C. Usual care (n=133) Treatment protocol: 10 weekly treatments, with first visits lasting 75 to 90 minutes and followup visits lasting 50 to 60 minutes	A. vs. B. vs. C 46 vs. 47 vs. 48 Mean age 66% vs. 65% vs. 62% female 86% vs. 87% vs. 86% white LBP Bothersomeness, VAS (0-10): 5.6 vs. 5.6 vs. 5.8 Disability, RDQ: 10.1 vs. 11.6 vs. 10.5	A. vs. B. LBP bothersomeness, VAS (0-10) mean change from baseline (10 weeks): A. vs. C: -1.4 (-1.9 to - 0.8) B vs. C: -1.7 (-2.2 to - 1.2) A. vs. B.: 0.3 (-0.2 to 0.8) P<0.05 but not reported separately	Disability, RDQ mean change from baseline (10 weeks): A. vs. C: -2.5 (-3.5 to -1.4) B vs. C: -2.9 (-4.0 to -1.8) A. vs. B.: 0.5 (-0.5 to 1.5) P<0.05 but not reported separately
Eghbali, 2012 (79) 2 weeks Chronic (Iran)	A: Reflexology (n=25) B: Nonspecific massage (n=25) 40 minute sessions three times a week for two weeks	A vs. B Mean age: 42 vs. 39 years Female: 40% (overall) Pain: 5.0 vs. 5.2 Function: Not reported	A vs. B Pain, VAS (0-10), mean change from baseline to 2 weeks: -2.28 vs. -1.36 (p<0.001)	Not reported
Kong, 2012 (72)	A: Chinese massage with herbal ointment (n=55)	A. vs. B. Mean age 21 vs.	A. vs. B. Pain mean change from	Disability not reported C-SFMPQ scores favored A. vs. B.

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
2 months Acute and chronic Good	B: Standard massage (n=55) Treatment protocol: 2 30 minute sessions per week x 4 weeks	20 (male athletes) 26/55 vs. 27/55 female Race not reported (Shanghai) Pain, 5.4 vs. 5.4 Disability, not reported	baseline (0-10 VAS): (-0.64 points [95% CI -1.04 to -0.24]; p=0.002	
Kumnerddee, 2009 (80) 10 days Chronic (Thailand)	A: Thai massage (n=8) B: Acupuncture (n=9) Treatment protocol: 5 sessions every 2 to 3 days over 10 days, massage sessions were 1 hour in duration	A vs. B Mean age: 26 vs. 29 years All male Race not reported (Thailand) Pain, VAS (0-10): 4.56 vs. 4.19 McGill score: 6.13 vs. 15.78	A vs. B Pain, VAS (0-10), mean change from baseline: 2.41 (1.73 to 5.74) vs. 2.41 (0.32 to 4.50)	A vs. B McGill score, mean change from baseline: 13.67 (6.91 to 20.42) vs. 5.88 (1.01 to 10.74)
Little, 2008 (29) 52 weeks Subacute, chronic Good	A. Exercise + 24 lessons in Alexander technique (n=71) B. Exercise + 6 lessons in Alexander technique (n=71) C. Exercise + massage (n=72) D. Exercise (n=72) E. 24 lessons in Alexander technique (n=73) F. 6 lessons in Alexander technique (n=73) G. Massage (n=75) H. Usual care (n=72)	A. vs. B. vs. C. vs. D. vs. E. vs. F. vs. G. vs. H. Mean age: 46 vs. 46 vs. 45 vs. 45 vs. 45 vs. 46 years Female sex: 73% vs. 78% vs. 63% vs. 64% vs. 68% vs. 71%	A. vs. B. vs. C. vs. D. vs. E. vs. F. vs. G. vs. H. Number of days of pain in previous 4 months, difference vs. usual care: -20 (p=0.001) vs. -13 (p=0.031) vs. -11 vs. -13 (p=0.001) vs. -13 (p=0.034) vs. -8 vs. 0 (ref)	A. vs. B. vs. C. vs. D. vs. E. vs. F. vs. G. vs. H. Roland disability score, difference versus usual care: -4.22 (p=0.002) vs. -2.98 (p=0.002) vs. -2.37 (p=0.015) vs. -1.65 vs. -4.14 (p<0.001) vs. -1.44 vs. -0.45 vs. 0 (ref) SF-36 PCS, difference vs. usual care: 9.43 (p=0.015) vs. 8.53 (p=0.029) vs. 3.63 vs. -2.08 vs. 11.83 (p=0.002) vs. 2.04 vs. -1.45 vs. 0 (ref) SF-36 MCS, difference vs. usual care: 4.99 vs. 0.64 vs. 2.73 vs. 0.72 vs. 3.74 vs. 4.10 vs. -2.11 vs. 0 (ref)
Quinn, 2008 (81) 18 weeks Chronic	A: Reflexology to specific reflex point on the feet (n=7) B: Sham massage (simple foot massage) (n=8)	A vs. B Median age: 42 vs. 45 years Female: 86% vs. 50%	A vs. B Pain, VAS (median, 0-10): 4.7 vs. 3.4 at baseline, 3.1 vs. 3.9 at 6 weeks, 2.1 vs. 4.1 at 12 weeks, 2.2 vs.	A vs. B RDQ, median: 5 vs. 7.5 at baseline, 6 vs. 5 at 6 weeks, 4 vs. 4.5 at 12 weeks, 4 vs. 3.5 at 18 weeks McGill pain questionnaire (0-77), median: 24

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
	Treatment protocol: 6 weekly 40 min sessions	Baseline pain (0-10 VAS): 4.7 vs. 3.4 Baselien RDQ: 5 vs. 7.5	3.2 at 18 weeks	vs. 19 at baseline, 12 vs. 11.5 at 6 weeks, 11 vs. 6.5 at 12 weeks, 6 vs. 7.5 at 18 weeks SF-36 Physical Component Summary: 38 vs. 43 at baseline, 46 vs. 44 at 6 weeks, 48 vs. 52 at 12 weeks, 48 vs. 52 at 18 weeks SF-36 Mental Component Summary: 38 vs. 43 at baseline, 46 vs. 44 at 6 weeks, 48 vs. 52 at 12 weeks, 48 vs. 52 at 18 weeks
Romanowski, 2012 (73) 10 days FU to 48 weeks Chronic Poor	A. Therapeutic massage (n=13) B. Deep tissue massage (n=13) Treatment protocol: 10 daily 30 min sessions	A. vs. B. Not described except to say there were no differences in age and sex	A. vs. B. Pain, VAS (0-10), mean change from baseline: 13.54 ± 7.75 vs. 4.92 ± 13.55 p<0.001	A. vs. B. ODI, mean change from baseline: 9.46 ± 11.22 vs. 16.38 ± 11.68, p<0.001
Sritoomma, 2014 (74) 15 weeks Chronic Fair	A. Swedish massage with ginger oil (n=70) B. Thai massage (n=70) Treatment protocol: 10 30 min sessions over a 5 week period	A. vs. B. Mean age not described (60 and older) 77% vs. 83% female Race not described (Thailand) Pain, VAS (0-10): 66.66 vs. 63.27 Disability, ODI: 26.9 vs. 29.5	A. vs. B. Pain, VAS (0-10) mean change from baseline: -6.37 (-12.58, -0.17) 0.044 (15 weeks)	A. vs. B. RDQ, mean change from baseline: - -3.66 (-7.17 to -0.14), p=0.042
Yoon, 2012 (82) 2 weeks Chronic (Korea)	A: Deep cross-friction massage (roptrotherapy) with the HT-bar (n=12), 20 minutes a day, 3 days a week for 2 weeks B: Transcutaneous electrical nerve stimulation (n=10), 20 minutes a day, 5 days a week for 2 weeks	A vs. B Mean age: 50 vs. 53 years Female: 58% vs. 60% Pain: 56.7 vs. 55.6 ODI: 34 vs. 30	A vs. B Pain (0-100 NRS): 56.7 vs. 55.6 at baseline (p=0.72), 23 vs. 34 at 2 weeks (p=0.07)	A vs. B ODI: 34 vs. 30 at baseline (p=0.31), 14 vs. 21 at 2 weeks (p=0.23) RDQ: 7.50 vs. 7.30 at baseline (p=0.77), 2.33 vs. 2.80 at 2 weeks (p=0.87)
Zhang, 2015 (78)	A. Chinese massage + core stabilization exercises (n=46)	A. vs. B. Mean age: 48.71	A. vs. B. VAS (0-100), 2 weeks:	A. vs. B. ODI, 2 weeks: 21.58±6.34 vs. 23.41±7.43,

**Author, Year
Duration of
Followup
LBP Duration
Quality**

Intervention and Duration of Treatment

Population

Pain Outcomes

Other Outcomes

1 year Duration of LBP: not reported Fair	B. Chinese massage only (n=46)	vs. 51.62 years Female: 37% vs. 33% Baseline pain: not reported Baseline function: not reported Duration of pain: ≥12 weeks: 43% vs. 37%	3.88±1.31 vs. 4.12±1.33, p>0.05 VAS (0-100), 8 weeks: 1.46±0.76 vs. 2.85±1.58, p<0.05	p>0.05 ODI, 8 weeks: 13.20±2.42 vs. 18.39±3.67, p<0.05
Zheng, 2012 (75) 3 weeks, Subacute, chronic Fair	A. Massage + traction (n=32) B. Traction alone (n=32)	A. vs. B. Mean age: 43 vs. 42 years Female: 44% vs. 50% Baseline pain: not reported Baseline function: not reported	A. vs. B. Mean difference in pain VAS (0-10) 1.9±0.9 vs. 1.4±0.8 p<0.05	
Balthazard, 2012 (88) 6 months Chronic Fair	A. HVLA + 5-10 min active exercises (n=22) B. Detuned ultrasound (sham) + 5-10 min active exercises (n=20) Treatment protocol: 8 sessions over 4-8 weeks (unclear duration)	A. vs. B. Mean age 44 vs. 42 years 36% vs. 30% female Race not reported Pain VAS (0-10) 53 vs. 65 ODI: 30 vs. 32	A. vs. B. Pain, VAS (0-10)-pain mean group difference: -1.24; 95% CI: -2.37 to - 0.30; p=0.032, statistically not significant at the 0.025 level.	A. vs. B. ODI mean group difference: -7.14; 95% CI: - 12.8 to - 1.52; p=0.013
Bicalho, 2010 (89) Immediate Chronic Fair	A. HVLA (n=20) B. Control (side lying) (n=20) Treatment protocol: single session	A. vs. B. Mean age 30 vs. 27 ODI: 14.6 vs. 16.6 Race not reported (Brazil)	A. vs. B. Pain VAS mean group difference (0-100): -11 vs. -2.2, no CI provided, p=0.04)	A. vs. B. Finger to floor, EMG flex-ext reported (favored SMT), ODI measured but not reported
Bronfort, 2004 (100) 52 weeks	A. Chiropractic (n=11) B. Epidural steroid injection (n=11) C. Self-care education (n=10)	A. vs. B. vs. C Mean Age: 44 vs. 52 vs. 52	Results were combined; no group specific results were reported.	Results were combined; no group specific results were reported. 3 weeks vs. 12 weeks vs. 52 weeks:

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Subacute or Chronic Poor		Female=45% v 36% v 50%	3 weeks vs. 12 weeks vs. 52 weeks: Leg Pain: 1.8 vs. 2.9 vs. 2.3 Low back pain: 0.9 vs. 1.7 vs. 1.9	RDQ: 13.7 vs. 22.7 vs. 19.6 Oswestry disability questionnaire: 11 vs. 22.9 vs. 15.6
Bronfort, 2011 (18) 52 weeks Chronic Good	A. Supervised exercise therapy for 12 weeks (n=100) B. Chiropractic spinal manipulation for 12 weeks (n=100) C. Home exercise and advice for 12 weeks (n=101)	A. vs. B. vs. C Mean age: 44.5 vs. 45.2 vs. 45.6 years Female: 57% vs. 66% vs. 58% Mean pain severity score (0-10): 5.1 vs. 5.4 vs. 5.2 Roland-Morris disability score (0- 23): 8.4 vs. 8.7 vs. 8.7		Only significant between-group differences in patient-reported outcomes were for satisfaction (favoring A, p<0.01 at 12 weeks and p<0.001 at 52 weeks) Overall treatment effect was significant for endurance (p<0.05) and strength (p<0.05) but not range of motion (also favoring A).
Bronfort, 2014 (99) 52 weeks Acute, subacute, chronic Fair	A. SMT plus HEA (home exercise with advice) (n=96) B. HEA (n=96)	A. vs. B. Mean age: 57 vs. 58 years Female: 59% vs. 68% Mean NRS (0-10) leg pain: 5.4 vs. 5.4 Mean RDQ: 10.2 vs. 10.2	A. vs. B. Leg Pain 12 weeks: -1.0 (- 1.9 to -0.2), p=0.008 Leg Pain 52 weeks: -0.7 (- 1.5 to 0.2), p=0.15 LBP 12 weeks: -0.9 (-1.6 to -0.3), p=0.005 LBP 52 weeks: -0.3 (-1.0 to 0.4) p=0.4	
Burton, 2000 (101) 12 months Chronic Poor	A. Osteopathic manipulation (15 min treatment sessions over 12 weeks) (n=20) B. Chemonucleolysis (control) (n=20)	Mean Age 42 53% female Mean duration of symptoms: 30 weeks vs. 32 weeks	A. vs. B. Leg Pain, baseline: 4 vs. 3.7; 2 weeks: 3.2 vs. 3.3; 6 weeks: 2.7 vs. 2.7; 12 months: 2.1 vs. 2.3 Back pain, baseline: 3.8 vs. 4.1; 2 weeks: 3.2 vs. 4; 6 weeks: 2.7 vs. 3.6; 12 months: 2.3 vs. 2.9	RDQ, baseline: 11.9 vs. 12; 2 weeks: 10.2 vs. 13.9; 6 weeks: 7.8 vs. 11; 12 months: 5.9 vs. 7.3
Cecchi, 2010 (90)	A. Back school (n=70) B. Physical therapy (n=70)	A. vs. B. vs. C. Mean age 58 vs.	A. vs. B. vs. C. Mean differences not	A. vs. B. vs. C.

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
12 months Chronic Fair	C. SMT (n=70) Treatment protocol: Back school and individual physical therapy: 15 1-hour-sessions for 3 weeks. SMT: 4-6 20 min sessions once a week	61 vs. 58 49% vs. 43% vs. 48% female Race not reported (Italy) Pain, (mean): 2 vs. 2 vs. 2.2 RDQ (0-24) (mean): 9.5 vs. 9.7 vs. 8.5 (sick leave due to LBP higher in A. vs. B. A. vs. B and C – p=0.001)	reported – will need to calculate Back Pain 12 month mean change from baseline (0.7 vs. 0.4 vs. 1.5) C improved to greater degree than B or A at 12 months in terms of pain (but small, clinically insignificant)	RDQ mean (SD) reduction from baseline to 12 months: 4.2+/- 4.8 vs. 4.0+/-5.1 vs. 5.9+/-4.6 C improved to greater degree than B or A at 12 months in terms of disability (but small, clinically insignificant)
De Oliveira, 2013 (91) immediate Chronic Good	A: HVLA – region specific (n=74) B: HVLA nonspecific (n=74) Treatment protocol: single treatment	A. vs. B. Mean age 46 vs. 46 80% vs. 68% female Race not reported Pain, NRS 6.1 vs. 6.0 Disability, RDQ: 11.3 vs. 9.3	A. vs. B. Pain, intensity mean group difference: 0.50 (-0.10 to 1.10), p=0.10	A. vs. B. Pressure pain thresholds measured, no difference between groups
Goertz, 2013 (95) 4 weeks Acute Fair	A: HVLA + standard medical care (n=45) B: Standard medical care (n=46) Treatment protocol: 2 visits weekly x 4 weeks	A. vs. B. Mean age 25 vs. 26 15% vs. 14% female 73% vs. 52% White, more missing in SMC Pain, NPRS 5.8 vs. 5.8 Disability, RDQ: 11 vs. 12.7	A. vs. B. Pain, intensity (NRS 0-10) mean group difference: 1.2 (0.2, 2.3) p=0.02	A. vs. B. Disability (RDQ): 4.0 (1.3, 6.7), p=0.004
Haas, 2014 (92) 1 year	A: Massage (n=100) B: Massage + 6 SMT (n=100) C: Massage + 12 SMT (n=100)	A. vs. B. vs. C. vs. D. Mean age 41 vs.	A. vs. D. Pain intensity, percentage responders (>50%) at 52	Disability score calculated, but unclear what measure

**Author, Year
Duration of
Followup
LBP Duration
Quality**

Intervention and Duration of Treatment

Population

Pain Outcomes

Other Outcomes

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Chronic Good	D. Massage + 18 SMT (n=100) Treatment protocol: 15 min sessions (18 total, unclear duration); 5 min hot pack, 5 min SMT or massage + 5 min sham ultrasound	41 vs. 42 vs. 41 49% vs. 49% vs. 49% vs. 52% female Nonwhite: 14% vs. 18% vs. 11% vs. 16% Pain, VAS (0-100) 52.2 vs. 51.0 vs. 51.6 vs. 51.5	weeks 10.6 (-3.2, 24.4), NS NS differences in A. vs. B., A. vs. C Only sig diff in 12 week A. vs. C. 21.1 (7.7, 34.6), p<0.025	
Spinal Manipulation Trials				
Paatelma, 2008 (98) 1 year Acute to chronic Fair	A. SMT (n=45) B. McKenzie (n=52), C. "advice only to be active" (n=37) Treatment protocol: A and B: 3-7 sessions (mean 6) C. one 45-60 min session	A. vs. B. vs. C. Mean age 44 vs. 44 vs. 44 42% vs. 29% vs. 35% female Race not reported (Finland) Pain, VAS (0-10) (median): 20 vs. 16 vs. 16 RDQ (0-24) (median): 9 vs. 9 vs. 8	A. vs. C. (12 months) Pain, intensity (VAS 0-10) mean group difference: -4 (-17 to 9) p=0.714 B vs. C Pain, intensity (VAS 0-10) mean group difference: - 10 (-23 to 2) p=0.144	A. vs. C. (12 months) Disability (RDQ): -3 (-6 to 0) p=0.068 B vs. C Disability (RDQ): -3 (-6 to 0) 0.028
Petersen, 2011 (93) 12 months Chronic Good	A. McKenzie exercise (n=175) B. SMT (n=175) Treatment Protocol: Max 15 sessions over 12 weeks (variable)	A. vs. B. Mean age 38 vs. 37 59% vs. 53% female Race not reported (Denmark) Pain (3 0-10 scales), 30/60 vs. 29/30 Disability, RDQ: 13 vs. 13	A. vs. B. Pain, intensity mean group difference: 2.8 (- 0.2 to 5.8) p=0.063 (12 months)	A. vs. B. Disability (RDQ): 1.5 (0.2 to 2.9) p=0.030 (12 months, favoring A)
Santilli, 2006 (102) 180 days	A. Active manipulation 5 days/week (n=53) B. Control (simulated manipulation) (n=49)	Mean age <40 Female 30% vs. 45%	Patients with reduction of local pain: 98% vs. 94% (NS)	NS difference between SF-36 results

Author, Year Duration of Followup LBP Duration Quality	Intervention and Duration of Treatment	Population	Pain Outcomes	Other Outcomes
Acute Good		Pain 6.4 vs. 6.4 Radiating Pain 5.3 vs. 5.1	Patients with reduction of radiating pain 100% vs. 83% (p<0.01)	
Schneider, 2015 (96) 6 months Acute, subacute Good	A. Manual thrust SMT (n=37) B. Mechanical assisted SMT (n=35) C. Usual care (n=35)	A. vs. B. vs. C. Mean age: 41 vs. 41 vs. 40 years Female: not reported Pain: 5.7 vs. 5.5 vs. 6.0 ODI: 33.9 vs. 33.1 vs. 34.6	A. vs. B. vs. C. adjusted group differences, mean (95% CI) Pain: -1.2 (-3.2 to 0.7) vs. -0.9 (-2.9 to 1.1) vs. 0.3 (-1.6 to 2.3)	A. vs. B. vs. C. adjusted group differences, mean (95% CI) ODI: 0.4 (-10.2 to 11.0) vs. 1.4 (-9.1 to 12.0) vs. 1.0 (-9.6 to 11.6)
Senna, 2011 (94) 10 months Chronic Fair	A. sham SMT (12 sessions over 1 month) (n=40) B. SMT (12 sessions over 1 month) (n=27) C. SMT (12 sessions over 1 month + every 2 weeks x 9 months) (n=27) Treatment protocol: 12 sessions over 1 month for initial treatments	A. vs. B. vs. C. Mean age 42 vs. 40 vs. 42 24% vs. 27% vs. 24% female Race not reported (Egypt) Pain, VAS (0-10) 41 vs. 42 vs. 43 ODI: 38 vs. 39 vs. 40	A. vs. B. vs. C. Pain, intensity mean group difference: A. vs. B. Unadjusted mean difference in VAS (0-10) at 1 month 4; at 10 months 0 A. vs. C. Unadjusted mean difference at 1 month 6, at 10 months 17 Results not reported as group mean differences – will need to calculate these; overall B and C improved to similar degree compared with A at 1 month, group C maintained the improvement through 10 months whereas B returned to baseline for both pain and function	
Von Heymann, 2013 (97) 12 weeks Acute Fair	A. SMT and placebo-diclofenac (n=37) B. Sham SMT and diclofenac (n=38) C. Sham SMT and placebo diclofenac. (n=25)	A. vs. B. vs. C. Mean age 34 vs. 38 vs. 39 36% vs. 38% vs. 46% female	A. vs. B. vs. C. (only reported to 9 days) Pain VAS (0-10) – unable to calculate group mean	A. vs. B. vs. C. A. vs. B.: Unadjusted mean difference in RDQ at 12 weeks: 3.0 RDQ - unable to calculate group mean differences based on the way presented

**Author, Year
Duration of
Followup
LBP Duration
Quality**

Intervention and Duration of Treatment

Population

Pain Outcomes

Other Outcomes

Race not reported
(Germany)
Pain, VAS (0-10)
41 vs. 42 vs. 43
ODI: 38 vs. 39 vs.
40

differences based on the
way presented (graphs)

And only A. vs. B. was
presented, not A. vs. B.
vs. C

(graphs)

BDI=Beck Depression Inventory; CI=confidence interval; CSQ=Coping Strategies Questionnaire; DASS=Depression Anxiety Stress Scales; EQ-5D=EuroQOL 5-dimension; EMG= Electromyography; GA=graded activity; GAD-2=Generalized Anxiety Disorder Scale; GPE=Global Perceived Effect scales; GX=graded exposure; HEA=home exercise with advice; HRQOL=health related quality of life; HSCL-25=Hopkins Symptom Checklist; HVLA=high velocity low amplitude; IPQ=Illness Perception Questionnaire; IQR=interquartile range; LBP=low back pain; LBPRS=Low Back Pair Rating Scale; MCE=motor control exercises; MPQ=McGill Pain Questionnaire; NNT=number needed to treat; NR=not reported; NRS=numeric rating scale; NS=non-significant; NSAIDS= Nonsteroidal anti-inflammatory drug; ODI=Oswestry Disability Index; OR=odds ratio; PCS=Pain Catastrophizing Scale; PDI= Pain Disability Index; PHQ-8=Patient Health Questionnaire; PMR=periodized musculoskeletal rehabilitation; PSEQ=Pain Self-Efficacy Questionnaire; PSFS=Patient-Specific Functional Scale; QBPDS=Quebec Back Pain Disability Scale; RDQ=Roland Morris Disability Questionnaire; RR=relative risk; SD=standard deviation; SF-12 PCS=SF-12 physical component scale; SF-12 MCS=SF-12 mental component scale; SFMPQ=Short Form McGill Pain Questionnaire; SMD=standard mean difference; SMT=spinal manipulation therapy; TBC=treatment-based classification; TSK=Tampa Scale for Kinesiophobia; VAS=visual analog scale; WHOQOL-BREF=World Health Organization Quality of Life- brief version

*SMD calculated from means and SD based on sample before attrition

^aΔ VAS for bothersomeness (at the end of treatments)=absolute value of [VAS for bothersomeness (baseline) – VAS for bothersomeness (end of treatments)] /VAS for bothersomeness (baseline) significances by 2-sample t test.

Supplement Table 3. Quality assessment of systematic reviews

Author, Year	"A priori" design provided?	Duplicate study selection and data abstraction? a. Study selection b. Data abstraction	Comprehensive literature search performed?	Non-English language studies considered for inclusion?	Conducted searches for unpublished (gray) literature?	List of included studies provided?	List of excluded studies provided with reasons?	Characteristics of the included studies provided?
Yoga								
Cramer, 2013 (37)	Yes	a. Not stated explicitly b. Yes	Yes	Yes	No	Yes	Yes	Yes
Psychological Therapies								

Author, Year	"A priori" design provided?	Duplicate study selection and data abstraction? a. Study selection b. Data abstraction	Comprehensive literature search performed?	Non-English language studies considered for inclusion?	Conducted searches for unpublished (gray) literature?	List of included studies provided?	List of excluded studies provided with reasons?	Characteristics of the included studies provided?
Henschke, 2010 (48)	Yes	a. Yes b. Yes	Yes	Yes	Unclear	Yes	Yes	Yes
Multidisciplinary rehabilitation								
Kamper, 2014 (54)	Yes	a. Yes b. Yes	Yes	Yes	No	Yes	No	Yes
Acupuncture								
Lam, 2013 (59)	Unclear	a. Yes b. Yes	Yes	Yes	No	Yes	No	Yes
Lee, 2013 (58)	Unclear	a. Yes b. Yes	Yes	Yes	Yes	Yes	No	Yes
Massage								
Furlan, 2010 (70)	Yes	a. Yes b. Yes	Yes	Yes	Yes	Yes	Yes	Yes
Spinal Manipulation								
Rubinstein, 2011 (87)	Yes	a. Yes b. Yes	Yes	Yes	Yes	Yes	Yes	Yes
Rubinstein, 2012 (86)	Yes	a. Yes b. Yes	Yes	Unclear	Yes, but excluded from analysis	Yes	Yes	Yes

Author, Year	Scientific quality of included studies: a. Assessed? b. Documented?	Sensitivity analyses or stratified analyses conducted according to study quality?	Study conclusions supported by the evidence? (Was study quality considered in the synthesis?)	Conflict of interest stated? a) Systematic Review b) Individual Studies	Multidisciplinary systematic review team?	Quality Rating
Yoga						
Cramer, 2013 (37)	a. 2009 Updated Method Guidelines for Systematic Reviews in the Cochrane Back Review Group b. Yes	Yes; high vs. low risk of bias; if heterogeneity	Study quality considered	a. Systematic review: Yes b. Individual studies: No	Unclear	Good
Psychological Therapies						
Henschke, 2010 (48)	a. Yes b. Yes	No	Yes (yes)	a. Yes b. No	Yes	Good
Multidisciplinary rehabilitation						
Kamper, 2014 (54)	a. Yes b. Yes	Yes	Yes	a. Yes b. No	Yes	Good
Acupuncture						
Lam, 2013 (59)	a. Yes b. Yes	No	Unclear	a. Yes b. No	No	Fair
Lee, 2013 (58)	a. Yes b. Yes	Yes	Yes	a. Yes b. No	No	Fair
Massage						
Furlan, 2010 (70)	a. Yes b. Yes	Yes	Yes	a. Yes b. No	Yes	Good
Spinal Manipulation						
Rubinstein, 2011 (87)	a. Yes b. Yes	Yes	Yes	a. Yes b. Yes	Yes	Good

Author, Year	Scientific quality of included studies: a. Assessed? b. Documented?	Sensitivity analyses or stratified analyses conducted according to study quality?	Study conclusions supported by the evidence? (Was study quality considered in the synthesis?)	Conflict of interest stated? a) Systematic Review b) Individual Studies	Multidisciplinary systematic review team?	Quality Rating
Rubinstein, 2012 (86)	a. Yes b. Yes	Yes	Yes	a. Yes b. Yes	Yes	Good

Supplement Table 4. Quality Ratings of Randomized Trials

Author, Year	Randomization	Concealed Treatment Allocation	Baseline Group Similarity	Patient Blinded	Care Provider Blinded	Outcome Assessor / Data Analyst Blinded	Cointerventions Avoided or Similar	Compliance Acceptable in All Groups
Exercise Trials								
Albaladejo, 2010 (16)	Yes	Yes	Yes	No	No	Yes	Unclear	Unclear
Albert, 2012 (17)	Yes	No	Yes	No	No	Yes	Unclear	Unclear
Bronfort, 2011 (18)	Yes	Yes	Yes	No	No	Yes	Unclear	Yes
Garcia, 2013 (19)	Yes	Yes	Yes	No	No	Yes	Unclear	Unclear
George, 2008 (20)	Yes	No	No	No	No	Yes	Unclear	Unclear
Hagen, 2010 (21)	Yes	No	Yes	No	No	Yes	Unclear	Unclear
Hartvigsen, 2010 (22)	Unclear	Yes	Yes	No	No	Unclear	Unclear	Unclear
Helmhout, 2008 (23)	Yes	Unclear	No	No	No	Unclear	Unclear	Unclear
Henchoz, 2010(24)	Unclear	Unclear	Yes	No	No	No	Unclear	No

Author, Year	Randomization	Concealed Treatment Allocation	Baseline Group Similarity	Patient Blinded	Care Provider Blinded	Outcome Assessor / Data Analyst Blinded	Cointerventions Avoided or Similar	Compliance Acceptable in All Groups
Hofstee, 2002 (25)	Yes	No	No	No	No	No	No	Unclear
Hurley, 2015 (26)	Yes	Yes	Yes	No	No	Yes	Unclear	No
Jensen, 2012 (27)	Yes	Yes	Yes	No	No	Yes	Unclear	Yes
Kell, 2011 (28)	Unclear	Unclear	Yes	No	No	Unclear	Unclear	Unclear
Little, 2008 (29)	Yes	Yes	Yes	No	No	Yes	Unclear	Yes
Macedo, 2012 (30)	Yes	Yes	Yes	No	No	Yes	Unclear	Unclear
Machado, 2010 (31)	Yes	Yes	Yes	No	No	Yes	Unclear	Yes
Pengel, 2007 (32)	Yes	Yes	Yes	Unclear	No	Yes	No	Unclear
Stankovic, 2012 (33)	Unclear	Unclear	Yes	No	No	No	Unclear	Unclear
Tai Chi Trials								
Hall, 2011 (35)	Yes	Yes	Yes	No	No	No	Unclear	Yes
Weifen, 2013 (36)	Unclear	Unclear	Yes	No	No	Yes	Unclear	Yes
Yoga Trials								
Aboagye, 2015 (44)	Unclear	Unclear	Yes	No	No	Yes	Unclear	No
Nambi, 2014 (38)	Yes	Unclear	Yes	No	Unclear	Unclear	Unclear	Unclear
Saper, 2013 (39)	Yes	Unclear	No	No	Unclear	Yes	Yes	No
Mindfulness-based stress reduction Trials								

Author, Year	Randomization	Concealed Treatment Allocation	Baseline Group Similarity	Patient Blinded	Care Provider Blinded	Outcome Assessor / Data Analyst Blinded	Cointerventions Avoided or Similar	Compliance Acceptable in All Groups
Cherkin, 2016 (45)	Yes	Yes	Yes	No	No	Yes	Unclear	No
Morone, 2016 (46)	Yes	Yes	Yes	No	No	Yes	Unclear	Yes
Morone, 2009 (47)	Yes	Yes	No	No	No	Unclear	Unclear	Yes
Psychological Therapies								
Kahn, 2014 (49)	Yes	Unclear	Unclear	Yes	No	Yes	Unclear	Unclear
Lamb 2010 (50), 2012 (51)	Yes	Yes	Yes	No	No	Yes	No	No
Siemonsma, 2013 (52)	Yes	Yes	Yes	No	No	Yes	Yes	Unclear
Vong, 2011 (53)	Yes	Unclear	Yes	Yes	No	Yes	Yes	No
Multidisciplinary rehabilitation trials								
Eisenberg, 2012 (55)	Yes	Unclear	Yes	No	No	Unclear	Unclear	Yes
Gatchel, 2003 (56)	Yes	Unclear	Unclear	No	No	Unclear	Unclear	Yes
Monticone, 2014 (57)	Yes	Unclear	Yes	Yes	No	Unclear	Yes	Yes
Acupuncture								
Cho, 2013 (60)	Yes	Yes	Yes	Yes	No	Yes	Unclear	Unclear
Hasegawa, 2014 (61)	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
Vas, 2012 (62)	Yes	Yes	Yes	Yes	No	Yes	Yes	Unclear

Author, Year	Randomization	Concealed Treatment Allocation	Baseline Group Similarity	Patient Blinded	Care Provider Blinded	Outcome Assessor / Data Analyst Blinded	Cointerventions Avoided or Similar	Compliance Acceptable in All Groups
Weiss, 2013 (63)	Unclear	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Unclear
Yeh, 2016 (69)	Unclear	Unclear	Yes	No	No	Unclear	Yes	Unclear
Yun, 2012 (64)	Yes	Yes	Yes	No	No	Yes	No	Unclear
Massage								
Ajimsha, 2014 (76)	Unclear	No	Yes	Yes	Yes	No	Yes	Unclear
Borges, 2014 (77)	Yes	Unclear	Yes	Yes	Unclear	No	Unclear	Unclear
Cherkin, 2011 (71)	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Eghbali, 2012 (79)	Unclear	No	Yes	Yes	Unclear	Unclear	Unclear	Unclear
Kong, 2012 (72)	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Kumnerddee, 2009 (80)	Unclear	Unclear	Yes	Yes	No	No	Yes	Unclear
Little, 2008 (29)	Yes	Yes	Yes	Yes	No	No	Yes	Unclear
Quinn, 2008 (81)	Yes	Yes	Yes	Yes	Unclear	No	Unclear	Unclear
Romanowski, 2012 (73)	Unclear	Unclear	Yes	Yes	Yes	No	Yes	Unclear
Sritooma, 2014 (74)	Yes	Unclear	Yes	Yes	No	No	No	Unclear
Yoon, 2012 (82)	Unclear	Unclear	Yes	Yes	No	No	Unclear	Yes

Author, Year	Randomization	Concealed Treatment Allocation	Baseline Group Similarity	Patient Blinded	Care Provider Blinded	Outcome Assessor / Data Analyst Blinded	Cointerventions Avoided or Similar	Compliance Acceptable in All Groups
Zhang, 2015 (78)	No	No	Yes	Yes	Unclear	No	Yes	Unclear
Zheng, 2012 (75)	Yes	Unknown	Yes	Yes	Unclear	Unclear	Unclear	Unclear
Spinal Manipulation								
Balthazard, 2012 (88)	Yes	Unclear	Yes	No	No	Unclear	Unclear	Unclear
Bicahlo, 2010 (89)	Yes	Unclear	Yes	No	No	Unclear	Unclear	Unclear
Bronfort, 2004 (100)	Yes	No	Yes	No	No	No	Unclear	Yes
Bronfort, 2011 (18)	Yes	Yes	Yes	No	Yes	Unclear	Unclear	Yes
Bronfort, 2014 (99)	Yes	Yes	Yes	No	No	No	Unclear	Yes
Burton, 2000 (101)	Yes	Yes	Yes	No	No	No	Unclear	Unclear
Cecchi, 2010 (90)	Yes	Unclear	No	No	No	Unclear	Unclear	Unclear
de Oliviera, 2013 (91)	Yes	Yes	Yes	Yes	No	Yes	Unclear	Unclear
Goertz, 2013 (95)	Yes	Yes	Yes	No	No	Yes	No	Yes
Haas, 2014 (92)	Yes	Yes	Yes	No	No	Yes	Unclear	Yes
Paatelma, 2008 (98)	Yes	Yes	Yes	No	No	Yes	Unclear	Unclear
Petersen, 2011 (93)	Yes	Yes	Yes	No	No	Yes	Yes	Unclear

Author, Year	Randomization	Concealed Treatment Allocation	Baseline Group Similarity	Patient Blinded	Care Provider Blinded	Outcome Assessor / Data Analyst Blinded	Cointerventions Avoided or Similar	Compliance Acceptable in All Groups
Santilli, 2006 (102)	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Unclear

Author, Year	Attrition Reported	Attrition Acceptable	Timing of Outcome Assessment in All Groups Similar	Intention-to-Treat Analysis	Is There a Registered or Published Protocol	Avoidance of Selective Outcomes Reporting	Quality Rating
Exercise Trials							
Albaladejo, 2010 (16)	Yes	Yes	Yes	Yes	Yes	Yes	Fair (but results reporting poor)
Albert, 2012 (17)	Yes	Yes	Yes	Yes	Yes	Unclear	Fair
Bronfort, 2011 (18)	Yes	Yes	Yes	Yes	Yes	Yes	Good
Garcia, 2013 (19)	Yes	Yes	Yes	Yes	Yes	Yes	Good
George, 2008 (20)	Yes	No	Yes	Yes	Yes	Unclear	Poor
Hagen, 2010 (21)	Yes	Yes	Yes	Yes	No	Unclear	Fair
Hartvigsen, 2010 (22)	Yes	Yes	Yes	Yes	Yes	Yes	Fair
Helmhout, 2008 (23)	Yes	Yes	Yes	Yes	Yes	Unclear	Poor
Henchoz, 2010(24)	Yes	Yes	Yes	Yes	No	Yes	Poor
Hofstee, 2002 (25)	Yes	Yes	Yes	Yes	No	Unclear	Poor
Hurley, 2015 (26)	Yes	No	Yes	Yes	Yes	Yes	Fair
Jensen, 2012 (27)	Yes	Yes	Yes	Yes	Yes	Yes	Good
Kell, 2011 (28)	No	Unclear	Yes	Unclear	No	Yes	Poor

Author, Year	Attrition Reported	Attrition Acceptable	Timing of Outcome Assessment in All Groups Similar	Intention-to-Treat Analysis	Is There a Registered or Published Protocol	Avoidance of Selective Outcomes Reporting	Quality Rating
Little, 2008 (29)	Yes	Yes	Yes	Unclear	Yes	Yes	Good
Macedo, 2012 (30)	Yes	Yes	Yes	Yes	Yes	Unclear	Fair
Machado, 2010 (31)	Yes	Yes	Yes	Yes	Yes	Unclear	Fair
Pengel, 2007 (32)	Yes	Yes	Yes	Yes	Yes	Unclear	Fair
Stankovic, 2012 (33)	Yes	No	Yes	Yes	Unclear	Unclear	Poor
Tai Chi Trials							
Hall, 2011 (35)	Yes	Yes	Yes	Yes	No	Yes	Fair
Weifen, 2013 (36)	No	Unclear	Yes	Unclear	No	Yes	Fair
Yoga Trials							
Aboagye, 2015 (44)	Yes	No	Yes	No	Unclear	No	Poor
Nambi, 2014 (38)	Yes	Yes	Yes	Yes	No	Unclear	Poor
Saper, 2013 (39)	Yes	Yes	Yes	Yes	Yes	Yes	Fair
Mindfulness-based stress reduction Trials							
Cherkin, 2016 (45)	Yes	Yes	Yes	Yes	Yes	Yes	Good
Morone, 2016 (46)	Yes	Yes	Yes	Yes	Yes	Yes	Good
Morone, 2009 (47)	Yes	Yes	Yes	Yes	Unclear	Yes	Poor
Psychological Therapies							

Author, Year	Attrition Reported	Attrition Acceptable	Timing of Outcome Assessment in All Groups Similar	Intention-to-Treat Analysis	Is There a Registered or Published Protocol	Avoidance of Selective Outcomes Reporting	Quality Rating
Kahn, 2014 (49)	Yes	Yes	Yes	Unclear	No	Yes	Fair
Lamb 2010 (50), 2012 (51)	Yes	Yes	Yes	No	Yes	Yes	Fair
Siemonsma, 2013 (52)	Yes	Yes	Yes	No	Yes	Yes	Fair
Vong, 2011 (53)	Yes	Yes	Yes	No	No	Yes	Fair
Multidisciplinary rehabilitation trials							
Eisenberg, 2012 (55)	Yes	Yes	Yes	Yes	No	Unclear	Good
Gatchel, 2003 (56)	No	NA	Yes	Unclear	Yes	Unclear	Fair
Monticone, 2014 (57)	Yes	Yes	Yes	Yes	No	Unclear	Good
Acupuncture							
Cho, 2013 (60)	Yes	Yes	Yes	Yes	Unclear	Yes	Good
Hasegawa, 2014 (61)	Yes	Yes	Yes	Yes	Yes	Yes	Good
Vas, 2012 (62)	Yes	Yes	Yes	Yes	Unclear	Yes	Good
Weiss, 2013 (63)	Yes	Yes	Yes	Yes	Unclear	Unclear	Poor
Yeh, 2016 (69)	Yes	No	Yes	Yes	Unclear	Unclear	Poor
Yun, 2012 (64)	Yes	Yes	Yes	Yes	Unclear	Yes	Fair
Massage							

Author, Year	Attrition Reported	Attrition Acceptable	Timing of Outcome Assessment in All Groups Similar	Intention-to-Treat Analysis	Is There a Registered or Published Protocol	Avoidance of Selective Outcomes Reporting	Quality Rating
Ajimsha, 2014 (76)	Yes	Yes	Yes	Yes	Unclear	Yes	Fair
Borges, 2014 (77)	Yes	Yes	Yes	Yes	Unclear	Yes	Fair
Cherkin, 2011 (71)	Yes	Yes	Yes	Yes	Unclear	Yes	Good
Eghbali, 2012 (79)	No	Unclear	Yes	Yes	Unclear	Yes	Poor
Kong, 2012 (72)	Yes	Yes	Yes	Yes	Unclear	Yes	Good
Kumnerddee, 2009 (80)	Yes	Yes	Yes	Yes	Unclear	Yes	Fair
Little, 2008 (29)	Yes	Yes	Yes	Unclear	Yes	Yes	Good
Quinn, 2008 (81)	Yes	Yes	Yes	Yes	Unclear	Yes	Fair
Romanowski, 2012 (73)	Yes	Yes	Yes	Yes	Unclear	Yes	Poor
Sritooma, 2014 (74)	Yes	Yes	Yes	Yes	Unclear	Yes	Fair
Yoon, 2012 (82)	Yes	Yes	Yes	Yes	Unclear	Yes	Fair
Zhang, 2015 (78)	Yes	Yes	Yes	Yes	Unclear	Yes	Fair
Zheng, 2012 (75)	Yes	Yes	Yes	Yes	Unclear	Yes	Fair
Spinal Manipulation							
Balthazard, 2012 (88)	Yes	Yes	Yes	Yes	Unclear	Yes	Fair

Author, Year	Attrition Reported	Attrition Acceptable	Timing of Outcome Assessment in All Groups Similar	Intention-to-Treat Analysis	Is There a Registered or Published Protocol	Avoidance of Selective Outcomes Reporting	Quality Rating
Bicahlo, 2010 (89)	Yes	Yes	Yes	Yes	Unclear	Yes	Fair
Bronfort, 2004 (100)	No	Unclear	Yes	Yes	Unclear	No	Poor
Bronfort, 2011 (18)	Yes	Yes	Yes	Yes	Unclear	Yes	Good
Bronfort, 2014 (99)	Yes	Yes	Yes	Yes	Yes	Yes	Fair
Burton, 2000 (101)	Yes	Unclear	Yes	Yes	Unclear	Yes	Poor
Cecchi, 2010 (90)	Yes	Yes	Yes	Yes	Unclear	Yes	Fair
de Oliveira, 2013 (91)	Yes	Yes	Yes	Yes	Yes	Yes	Good
Goertz, 2013 (95)	Yes	No	Yes	Yes	Yes	Yes	Fair
Haas, 2014 (92)	Yes	Yes	Yes	Yes	Unclear	Yes	Good
Paatelma, 2008 (98)	Yes	No	Yes	Yes	Unclear	Yes	Fair
Petersen, 2011 (93)	Yes	Yes	Yes	Yes	Unclear	Yes	Good
Santilli, 2006 (102)	Yes	Yes	Yes	Yes	Unclear	Yes	Good

Supplement Table 5. Nonpharmacologic treatments versus active comparators for acute or subacute low back pain

Intervention	Pain: Magnitude of Effect	Evidence	SOE	Function: Magnitude of Effect	Evidence	SOE
--------------	---------------------------	----------	-----	-------------------------------	----------	-----

Exercise vs. exercise	No clear differences	> 20 RCTs	Moderate	--	--	--
Spinal manipulation vs. other active interventions	No clear differences at 1 week, 1 month, 3- 6 months, 1 year	1 SR (3 RCTs)	Moderate	No clear differences	1 SR (3 RCTs)	Moderate
Spinal manipulation plus exercise or advice vs. exercise or advice alone	--	--	--	Small, favors spinal manipulation at 1 week	1 SR (4 RCTs)	Low
Spinal manipulation plus exercise or advice vs. exercise or advice alone	--	--	--	No clear differences at 1, 3 months	1 SR (3 RCTs)	Low

RCT=randomized controlled trial, SOE=strength of evidence, SR=systematic review

Supplement Table 6. Nonpharmacologic treatments versus active comparators for chronic low back pain

Pain Intervention	Pain: Magnitude of Effect	Evidence	SOE	Function: Magnitude of Effect	Evidence	SOE
Motor control exercise vs. general exercise (short term)	Small, favors MCE for short term	1 SR (6 RCTs)	Low	Small, favors MCE	1 SR (6 RCTs)	Low
Motor control exercise vs. general exercise (intermediate term)	Small, favors MCE for intermediate term	1 SR (3 RCTs)	Low	--	--	--
Motor control exercise vs. general exercise (longer term)	Small, favors MCE for longer term	1 SR (4 RCTs)	Low	Small, favors MCE	1 SR (3 RCTs)	Low
Motor control exercise vs. multimodal physical therapy (intermediate term)	Moderate, favors MCE	1 SR (4 RCTs)	low	Moderate, favors MCE	1 SR (3 RCTs)	Low
Motor control exercise plus exercise vs. exercise alone	No clear differences	2 RCTs	Low	--	--	--
Pilates vs. usual care plus physical activity	No effect to small effect, favors Pilates	7 RCTs	Low	No clear differences	7 RCTs	Low
Pilates vs. other exercise	No clear differences	3 RCTs	Low	No clear differences	3 RCTs	Low
Tai chi vs. other exercise	Moderate, favors tai chi	1 RCTs	Low	--	--	--
Yoga vs. exercise	Small, favors yoga	1 SR (5 RCTs)	Low	--	--	--
Mindfulness-based stress reduction vs. cognitive behavioral therapy	No clear differences	1 RCT	Low	No clear differences	1 RCT	Low
Psychological therapies vs. exercise or physical therapy	No clear differences	1 SR (6 RCTs)	Low	--	--	--
Psychological therapies vs. psychological therapies	No clear differences	10 RCTs	Moderate	No clear differences	10 RCTs	Moderate
Multidisciplinary rehabilitation vs. physical therapy (short term)	Small, favors multidisciplinary rehabilitation	1 SR (12 RCTs)	Moderate	Small, favors multidisciplinary rehabilitation	1 SR (13 RCTs)	Moderate
Multidisciplinary rehabilitation vs. physical therapy (long term)	Moderate, favors multidisciplinary rehabilitation	1 SR (9 RCTs)	Moderate	--	--	--
Multidisciplinary rehabilitation vs. physical therapy (long term)	Moderate, favors multidisciplinary rehabilitation	1 SR (10 RCTs)	Moderate	--	--	--
Spinal manipulation versus other active interventions (exercise, usual care, medications, massage)	No clear differences	1 SR (6 RCTs)	Moderate	No clear differences	1 SR (6 RCTs)	Moderate
Acupuncture vs. medications	Small, favors acupuncture	1 SR (3 RCTs)	Low	Small, favors acupuncture	1 SR (3 RCTs)	Low

MCE=motor control exercise, RCT=randomized controlled trial, SOE=strength of evidence, SR=systematic review