REVIEW ARTICLE

EFFECTIVENESS OF EXERCISE ON WORK DISABILITY IN PATIENTS WITH NON-ACUTE NON-SPECIFIC LOW BACK PAIN: SYSTEMATIC REVIEW AND META-ANALYSIS OF RANDOMIZED CONTROLLED TRIALS

Peter Oesch, MSc, PT[1,2], Jan Kool, PhD, PT[3], Kare Birger Hagen, PhD, PT[4] and Stefan Bachmann, MD[1,2,5]

From the [1]Research Department, [2]Department of Rheumatology, Rehabilitation Centre Valens, Valens, [3]Zurich University of Applied Science, School of Physiotherapy, Zurich, Switzerland, [4]University of Oslo, Institute of Health Sciences and Nursing, Oslo, Norway and [5] Department of Geriatrics, Inselspital, Bern University, and University of Bern, Switzerland

Objectives: To determine whether exercise is more effective than usual care to reduce work disability in patients with non-acute non-specific low back pain, and if so, to explore which type of exercise is most effective.

Methods: Systematic review and meta-analysis of randomized controlled trials investigating the effectiveness of exercise in non-acute non-specific low back pain, and reporting on work disability. Data sources: MEDLINE, EMBASE, PEDro, Cochrane Library databases, NIOSHTIC-2, and PsycINFO until August 2008. Work disability data were converted to odds ratios. Random effects meta-analyses were conducted.

Results: A total of 23 trials met the inclusion criteria, 20 of which were suitable for inclusion in meta-analysis allowing 17 comparisons of exercise interventions with usual care and 11 comparisons of 2 different exercise interventions. A statistically significant effect in favour of exercise on work disability was found in the long term (odds ratio (OR) = 0.66, 95% confidence interval (CI) 0.48–0.92) but not in the short (OR = 0.80, 95% CI 0.51–1.25) and intermediate term (OR = 0.78, 95% CI 0.45–1.34). Meta-regression indicated no significant effect of specific exercise characteristics.

Conclusion: Exercise interventions have a significant effect on work disability in patients with non-acute non-specific low back pain in the long term. No conclusions can be made regarding exercise types.

INTRODUCTION

Low back pain is the most prevalent of musculoskeletal conditions. It affects almost everyone during their lifetime and has become a major socioeconomic problem in western countries (1). Exercise is consistently recommended in modern treatment guidelines for non-acute non-specific low back pain (NSLBP)

defining return to work as the primary treatment goal (2, 3). Exercises applied in the treatment of patients with NSLBP encompass a wide variety of interventions and are applied with different rationales. The sports medicine approach applies exercise based on the principles of exercise physiology, and is used in functional restoration programs with the goal of restoring physical function and thereby enabling patients to return to work (4). Behavioural treatment programmes use exercise with the aim of modifying pain behaviour. Patients learn that it is safe to move, while restoring function by receiving continuous feedback and positive reinforcement (5).

Until the year 2000 no evidence was found for the effectiveness of specific exercises in the management of chronic low back pain. Abenhaim et al. (2) state: "it appears that the key to success is physical activity itself, i.e. activity of any form, rather than any specific activity". A later review revealed that individually designed stretching or strengthening exercises delivered with supervision may improve pain and function in chronic NSLBP. The authors recommended further testing with this multivariable model and further assessment with specific patient-level characteristics and exercise types (6). A systematic review of trials with positive outcomes on work disability revealed that all had significant cognitive behavioural components combined with intensive physical training. The authors, however, advised caution when interpreting this *post-hoc* analysis and recommended further investigation into the contribution of these exercise characteristics (7). Whereas additional reviews found limited evidence for the effectiveness of behavioural graded activity in improving absenteeism outcomes (8, 9), strong evidence has been found that exercise reduces work disability in patients with NSLBP (7, 10, 11). These reviews were based on studies published prior to 2004 that did not evaluate the effectiveness of different exercise characteristics.

Although new studies have been published in the meantime, the effect of specific exercise characteristics on work disability is still unclear; a more up-to-date review is required. The objective of this review is to use recent research results to determine whether exercise is more effective than usual care to reduce work disability in patients with non-acute NSLBP, and if so, to explore which type of exercise is most effective.

METHODS

Searching

The search strategy was based on the recommendations of the Cochrane Back Review Group (12). We searched MEDLINE, EMBASE, PEDro, the Cochrane Library, PsycINFO (2002–August 2008) and NIOSHTIC-2 (until August 2008). This search was combined with a previous search performed in December 2002 covering MEDLINE (1966–October 2002), EMBASE (1988–October 2002), PEDro (until December 2002), the Cochrane Library (2002, Issue 4) and PsycLIT (1984–December 2002) (10). References were checked for further trials.

Selection, validity assessment, and data abstraction

Studies were included if randomized controlled trials were performed, the primary diagnosis in all patients was non-acute NSLBP with a duration of at least 4 weeks, the experimental treatments used exercise alone or as a part of a multidisciplinary treatment, work disability was the primary outcome, and if at least 90% of the patients under treatment were available for the job market, in that they were either employed or unemployed but seeking work. Excluded were studies that did not report work disability, investigated the effect of treatments that did not contain any form of exercise such as respondent psychological interventions, included patients with thoracic pain, cervical pain or specific low back pain, such as nerve root compression, vertebral fracture, tumour, infection, inflammatory diseases, spondylolisthesis, spinal stenosis and definite instability, and studies that included pregnant women with low back pain.

Two authors (PO and JK) independently applied the admission criteria for the studies and assessed risk of bias. Disagreements were solved through discussion involving a third researcher (StB). Authors were contacted if the information regarding the eligibility of a trial, quality criteria, or work disability were unclear.

Study quality was assessed according to Juni et al. (13), who stated that the internal validity of a study was threatened by detection bias, attrition bias, selection bias, and performance bias. Thus, the following 3 criteria were rated as "met", "unclear" or "not met": Concealed allocation, blinding of the outcome assessor, and intention to treat analysis. Performance bias was not assessed as it is not strictly possible to blind the treatment provider and recipient in clinical trials investigating the effect of exercise to treatment allocation. The internal validity of the included studies were then evaluated on methodological overall assessment. Studies were classified as high-quality studies if 2 or 3 of the criteria were met, while studies were classified as of low quality if one or none of the criteria were met.

For each study, 2 of the authors (PO and StB) independently extracted data from all included studies and defined exercise characteristics. Four criteria designed by Hayden et al. (6) were used, namely programme design, delivery type, dose, and type. Additionally, 2 criteria proposed by Schonstein et al. (14) were used, namely work context and exercise administration within a cognitive behavioural approach. A further criterion was the setting in which exercise was applied (see Table I).

Quantitative data synthesis

Work-related outcomes were converted into odds ratios (OR) using the method described by Chinn (15) and Hasselblad & Hedges (16). This method is based on the fact that, when assuming logistic distributions and equal variances in the 2 treatment groups, the log OR corresponds to a constant multiplied by the standardized difference between means. The ''metan'' command for Stata statistical software (Stata Corporation, College Station, TX V10) was used to conduct DerSimonian and Laird random effects meta-analyses (17). To use all available means, we estimated missing standard deviations (SD) from other included studies. We assessed treatment effects at 3 different times of follow-up (short-term = closest to 4 weeks, intermediate-term = closest to 6 months, long-term = closest to 12 months). Between-trial heterogeneity was quantified using the I^2 statistic, which can be understood as the proportion of the total variation in estimated ORs that is due to between-trial heterogeneity rather than chance (18). The extent

Table I. Exercise intervention characteristics

Programme design (according to Hayden et al. (6))
- "Individually designed", in which the treating therapist completed a clinical history and physical examination and delivered an exercise programme specifically designed for the individual participant.
- "Standard design", in which a fixed exercise programme was delivered to all participants.

Delivery type (according to Hayden et al. (6))
- Home exercises: participants performed their exercises at home with no direct supervision by the therapist.
- Supervised exercises: participants performed their exercises either under 1-on-1 supervision or attended exercise therapy sessions with 2 or more participants.
- Exercise therapy programmes that included both types of delivery will be classified according to their main delivery type.

Exercise dose (hours of intervention time) (according to Hayden et al. (6))
- We will calculate the exercise dose from the exercise duration and the number of treatment sessions received. Home exercise dose will only be included in exercise dose calculation if the home training was controlled (i.e. by using an exercise diary or by follow-up visits). If the study adherence information is not reported, we will assume an adherence rate of 50%. We will then dichotomize the exercise interventions into high- and low-dose exercise.

Types of exercises (according to Hayden et al. (6))
- Strengthening exercises
- Stretching exercises
- Mobilizing or flexibility exercises
- Aerobic exercises
- Stabilization exercises
- Programmes that included different exercise types will be classified as mixed exercises.

Setting
- Inpatient
- Outpatient

Work context (proposed by Schonstein et al. (14))
- Exercises will be classified as work-related if these were specifically designed to restore work-related physical capacity.

Behavioural treatment approach (proposed by Schonstein et al. (14))
- Exercises will be considered as administered within a behavioural treatment approach if this was specifically stated or if at least 3 of the following behavioural treatment modalities were applied: positive reinforcement of healthy behaviours (i.e. reassurance that it is safe to move, encouragement for early return to work); goal-contingent instead of pain-contingent exercise administration (i.e. exercise intensity was progressively increased to pre-set goals despite pain provocation); patients were given self-responsibility for treatment; patient education about a multidimensional view of pain (i.e. explanation of pain mechanisms); pain-coping strategies were applied (i.e. relaxation techniques were a consistent feature of the exercise programme).

to which one or more study characteristics explained between-trial heterogeneity was explored using meta-regression. The following explanatory variables were considered according to an *a priori* statistical analysis plan: exercise design (individual vs standard care), dose (high- vs low-dose exercise), delivery type (home-based exercises vs supervised exercises), type (specific vs mixed), administration within a cognitive behavioural approach (yes/no), work context (yes/no), and setting (in- vs outpatient) in bivariate models. In addition, we assessed the effect of methodological quality (low vs high). For work disability we included the variables above in meta-regression models and conducted random effects meta-analyses within each subgroup. Differences between small and large trials were assessed using funnel plots (19).

Table II. *Study characteristics*

Study	Participants	Interventions (exercise characteristic classification)	Work disability outcomes and terms of follow-up	Results
Alaranta et al., 1994 (Finland) (20)	293 patients with NSLPB > 6 months, age 30–47 years	A: Akseli Programme: 3 weeks, 37 h/week of guided or self-controlled physical exercise, 5 h/week cognitive-behavioural disability management groups, no passive treatments: (ID, SE, HD, MxE, IP, WCy, Bty) B: Current national type of inpatient rehabilitation: 3 weeks, 15–20 h/week physical activity, large amount of passive treatments: (ID, SE, HD, MxE, IP, WCn, BTn)	Number of subjects with more than 30 sick days at one year FU	No difference regarding sick leave (A: 26%, B: 23%) after one year
Aure et al., 2003 (Norway) (21)	49 patients sick-listed 8–24 weeks due to NSLPB, age 20–60 years	A: Manual therapy and specific exercise: 16 session lasting 45 min including spinal manipulation, specific mobilization, and stretching techniques as well as 5 general and 6 specific localized exercises for spinal segments and pelvic girdle, individual home exercises: (ID, SE, LD, SaE, OP, WCn, BTn) B: Supervised exercise therapy: 16 sessions consisting of 10 min warm-up on an exercise bicycle and 35 min. exercise session for the trunk, and lower limb muscles, suited to the clinical findings, individual home exercises: (ID, SE, LD, MxE, OP, WCn, BTn)	Proportion of partly or fully sick-listed patients at 4 weeks, 6 and 12 months FU	Significant advantage at 4 weeks FU for A vs B (30% vs 57%), at 6 months (11% vs 62%), and at 12 months (19% vs 59%).
Bendix et al., 1995, 2000, 1998 (Denmark) (22–24)	132 patients with disabling NSLPB > 6 months, threatened job situation or out of work	A: Functional restoration: full-time intensive 3-week multidisciplinary programme for 39 h per week, including active physical and ergonomic training and psychological pain management, followed by one 6-h day weekly for the subsequent 3 weeks: (ID, SE, HD, MxE, OP, WCy, BTy) B: Active physical training: Two sessions a week for 6 weeks, consisting of 45 min aerobics class, 45 min progressive resistance training, and 1 h of theoretical back lessons every second session: (SD, SE, LD, MxE, OP, WCn, BTn) C: Psychological pain management: 2 sessions lasting 75 min, twice a week for 6 weeks., combining active physical training of 15 min warm-up exercises, and 30 min progressive resistance training with psychological pain management: (SD, SE, LD, MxE, OP, WCn, BTn)	Sick days (median, IQR) and work capability as judged by a physician at 4 and 13 months FU	Sick days: significant advantage after 4 months for A and B vs C. A: 25 (0–103), B: 13 (0–122), C: 122 (60–122), and after 13 months for A vs B and C and for B vs C. A: 52 (0–127), B: 100 (0–390), C: 295 (0–390). Significantly better work capability in A vs B and C after 4 months (A: 75%, B: 48%, C: 40%) and 2 years (A: 80%, B: 55%, C: 44%)
Bendix et al., 1998, 1996 (Denmark) (24, 25)	106 patients with > 6 months of disabling back pain, age 18–59 years	A: Functional restoration: full-time intensive 3-week multidisciplinary programme for 39 h per week, including active physical and ergonomic training and psychological pain management, followed by 1 6-hour day weekly for the subsequent 3 weeks: (ID, SE, HD, MxE, OP, WCy, BTy) B: Usual care	Sick days (median, IQR) and work capability as judged by a physician at 4 months and 2 years FU	Sick days: significant difference at 4 months FU. A: 10 (0–122), B: 122 (24.5–122). Work capability: significantly larger improvement at 4 months. A: from 27% to 64%, B: from 16% to 29%. No difference after 2 years. A: 52%, B: 51%
Bendix et al., 2000 (Denmark) (26)	138 patients, precarious work situation due to NSLPB	A: Functional restoration: full-time intensive 3-week multidisciplinary programme for 39 h per week, including active physical and ergonomic training and psychological pain management, followed by one 6-hour day weekly for the subsequent 2 weeks, and a third day after 2nd month: (ID, SE, HD, MxE, OP, WCy, BTy) B: Outpatient intensive physical training: aerobics and strengthening exercises for 1.5 h 3 times/week for 8 weeks: (SD, SE, HD, MxE, OP, WCn, BTn)	Sick days (median, IQR) and work capability as judged by a physician at one year FU. No statistical comparison between groups. Results disregard 32 of 138 randomized patients who did not finish treatment	Sick days: No difference at 1 year FU. A: median in 34/48 patients 5.5 (0–113), B: median in 40/50 patients 2.5 (0–301). Work capability: pre-post A: from 28 to 36/48, B: from 21 to 35/51
Hagen et al., 2000, 2003 (Norway) (27, 28)	457 patients with NSLPB and sick leave 8–12 weeks, age 18–60 years	A: Light mobilization programme: examination at spine clinic for 1 hour, information, advice to stay active and go on daily walks; individual instructions lasting 1–1.5 h on stretching and training at home by the physical therapist: (ID, HE, LD, ScE, OP, WCn, BTy) B: Usual care by GP	Days of sickness compensation (mean, 95% CI), numbers of workers returning to full-duty work at 3, 6, 12, 24, and 36 months FU	Significant fewer sick days at one year FU in A vs B (A: 95.5, 104.5, B: 133.7, 112). RTW: significant improvement after 6 months in A vs B (A: 61.2 %, B: 45.0 %) and 1 year (A: 68.4 %, B: 56.4 %). No significant difference at 3-year FU (A: 36.2%, B: 38.2%)

Study	Population	Intervention	Outcome	Results
Härkäpää et al., 1989, 1990 (Finland) (29, 30)	476 blue-collar workers, age 35–54 years, sick leave due to NSLPB during the last 2 years	A: Inpatient rehabilitation: groups of 6–8 patients, 3 weeks, 4 sessions Swedish back school, 15 back exercise sessions, 9 relaxation sessions, heat or electrotherapy prior to exercise, 2 structured group discussions, home programme, rehearsal after 1.5 years (2 weeks): (SE, MxE, IP, WCn, BTn) B: Outpatient treatment at the workplace or local health centre: 15 sessions in 2 months, groups of 6–8 patients, 4 sessions Swedish back school, 15 back exercise sessions, 9 relaxation sessions, heat or electrotherapy prior to exercise, 2 structured group discussions, home programme, rehearsal after 1.5 years (8 sessions): (SE, MxE, OP, WCn, BTn) C: Physician examination: no systematic treatment, written and oral instructions on back exercises and ergonomics.	Sick days at 15 months FU disregarding the first 7 days of each episode of sickness absence leading to a considerable underestimation of the days lost from work	Sick days after 1.5 years: A vs C no difference. (A: 5.5, 25, C: 7.5, 25). B vs C no difference (B: 5.8, 25, C: 7.5, 25.0). A vs B no difference (A: 5.5, 25.0, B: 5.8, 25.0)
Heymans et al., 2006 (Nether-lands) (31)	299 workers sick-listed 3 to 6 weeks because of NSLPB, age 18–65 years; able to complete Dutch questionnaires	A: High-intensity back school: 2 sessions of 1 hour per week during 8 weeks, supervised by a physiotherapist. Instruction of individual strength training exercises, work simulation and home exercises during the treatment period. A cognitive-behavioural therapy approach was used: (ID, SE, LD, SgE, OP, WCy, BTy) B: Low-intensity back school: 4 group sessions based on the Swedish model, once a week, 4 consecutive weeks. Sessions were divided into 30 min education and a 90-minute practical part, comprised of a standardized exercise programme consisting of strength training and home exercises: (SD, SE, LD, SgE, OP, WCn, BTn) C: Usual care: provided by occupational physician according to the Dutch guidelines for the occupational health management of patients with LBP.	Cumulative percentages of RTW during 6 month FU. Sick days (mean, SD personal communication) at 6 month FU	Faster RTW was observed in B compared with C and A with hazard ratios of 1.4 (p=0.06) and 1.3 (p=0.09) respectively. No statistically significant difference regarding sick days between the 3 groups at 6 month FU (A: 96.1, 63.3; B: 81.8, 55.6; C: 92.5, 65.5)
Hurri, 1989 (Finland) (32)	80 female employees with NSLPB > 12 months	A: Back school given by a physical therapist; 6 sessions of 1 hour, consisting of education and exercise, in groups of 11 patients. Two review sessions of 1 hour after 6 months: (SD, SE, LD, MxE, OP, WCn, BTn) B: Usual care by GP and provision of instruction material from back school.	Sick days due to LBP at 12 months FU (mean, SD personal communication)	No difference during the 1st year (A: 8.1, 26.9, B: 11.1, 26.6) and during the 2nd year (A: 9.0, 23.6; B: 9.5, 25.0)
Jousset et al., 2004 (France) (33)	86 workers, threatened job situation by LBP, age 18–50 years, no limited work contract	A: Functional restoration programme: 6 h per day, 5 days/week during 5 weeks; muscular warm-up and stretching techniques, aerobic exercises, progressively increasing strengthening exercises, work simulation, Balneology for 30 min per day. Psychological counselling if required and dietary advice: (ID, SE, HD, MxE, WCy, BTn) B: Active individual therapy: Three 1-hour sessions per week during 5 weeks, supervised by private practice physiotherapists, instruction of a 50 min active home exercise programme twice a week, and recommendation of sports activities: (ID, SE, HD, MxE, OP, WCn, BTn)	Sick days (mean, SD) at 6 months FU and RTW (%) within one week	Sick days: advantage of A vs B (A. 28.7, 44.6; B: 48.3, 66.0). RTW no difference between the 2 groups (A 69.2%, B 66.7%)
Kääpa et al., 2006 (Finland) (34)	120 women with daily or nearly daily LBP, age 22–57 years	A: Outpatient multidisciplinary rehabilitation: programme of 70 h duration conducted in groups of 6 to 8 patients during 8 weeks. The intervention comprised cognitive-behavioural treatment, back school education including occupational intervention, and a physical exercise programme for 2–3 h per day: (ID, SE, HD, MxE, OP, WCn, BTy) B: Individual physiotherapy: Ten 1-hour treatment sessions during 6–8 weeks. 30–40-min passive pain treatment and 15–20-min light active exercise per session, advice to progressively increase regular daily activities, recommendation of general physical training, instruction of a light home-exercise programme: (ID, SE, LD, MxE, OP, WCn, BTn)	Percentage of patients in different categories of sick leave (0 days, 1–30 days, > 30 days) at 6, 12 and 24 months FU	No statistically significant differences of A vs B at all terms of FU (1 year: 0 days = A: 61.5%, B: 56.4%; 1–30 days = A: 32.7%; B: 34.6%; > 30 days = A: 5.8%; B: 9.0%; 2 years: 0 days = A: 68%; B: 63.8%; 1–30 days = A: 20%; B: 25.6%; >30 days = A: 12%; B: 10.6%.

Study	Population	Intervention	Outcome measure	Results
Karjalainen et al., 2003, 2004 (Finland) (35, 36)	164 patients, age 25–60, working difficulties between 4 and 12 weeks	A: Mini-intervention: 60 min physician examination, information on the good prognosis of back pain, and disadvantages of rest, encouragement to stay active followed by a 90-minute physiotherapy session; appraising of daily back-straining activities and training of special movements required at work, and instruction of a daily exercise programme: (ID, HE, LD, MxE, OP, WCy, BTy) B: Work site visit: mini intervention plus work site visit C: Usual care by GP	Sick days (mean, SD estimated as 2/3 of mean) and percentage of patients on sick leave at 12 and 24 months FU	Significantly fewer sick days at 1 year FU in A vs C (A: 19, 12; C 41, 27) and at 2 year FU. Significantly fewer patients on sick leave at 1 year FU (A: 45%, C: 61%)
Kool et al., 2005, 2007 (Switzerland) (37, 38)	174 patients with non-acute LBP, out of work > 6 weeks, age 20–55 years	A: Function centred treatment: Inpatient rehabilitation of 3 weeks duration, 4 h per day, focusing on improved function not pain reduction. Work simulation, strength and endurance training, sports therapy, and self-exercises: (ID, SE, HD, MxE, IP, WCy, BTy) B: Pain-centred treatment: inpatient rehabilitation of 3 weeks, 2.5 h per day, focusing on pain reduction. Individually selected passive and active exercises, mini back school, passive pain modulating treatments, and relaxation techniques: (ID, SE, HD, MxE, IP, WCn, BTn)	Sick days (mean, SD personal communication) at 12 months FU	Advantage after 12 months for A vs B (A: 247, 137; B: 291, 115)
Lindström et al., 1992, 1995, 1992 (Sweden) (5, 39, 40)	103 blue-collar workers, sick-listed 6 weeks, no NSLPB sick-listing in the prior 12 weeks	A: Graded activity programme: treatment until RTW achieved, no predefined treatment duration. Functional capacity testing, workplace visit, back school, individual sub-maximal gradually increasing exercise programme to teach the patient that it is safe to move while restoring function: (ID, SE, HD, MxE, OP, WCn, BTy) B: Usual care by GP	Number of patients RTW at 6 and 12 weeks FU, and sick days (mean, SD) during 2nd year FU	Significantly larger proportion of patients RTW in A vs B after 6 weeks (A: 59%, B: 40%) and 12 weeks (A: 80%; B: 58%). Sick days: significant advantage during the 2nd follow-up year of A vs B (A: 60, 92; B: 98, 103.5)
Niemisto et al. 2003, 2005 (Finland) (41, 42)	204 patients with chronic LBP, age 24–64 years, Oswestry Disability Index score > 16%	A: Combined treatment: physician's consultation incl. explanation of clinical findings, instructions regarding posture and exercise, and an educational booklet combined with manipulative treatment, and instruction of pain-free stabilizing exercises 4 times in the course of 4 week: (ID, SE, LD, SaE, OP, WCn, BTn) B: Physician's consultation: 1 hour at entry and 5 months FU	Sick days (mean, SD) at 1 and 2 years FU	No significant differences of A vs B at 1 year FU (A: 13.9, 26.6; B: 18.5, 38.8) and 2 years FU (A: 12.3, 20.5; B: 14.8, 38.0)
Petersen et al., 2002, 2007 (Denmark) (43, 44)	260 people with LBP > 8 weeks, age 18–60 years	A: McKenzie treatment: examination by physiotherapist 1 hour, treatment sessions 30 min (maximum 15 sessions) consisting of self-mobilization with repeated movements, mobilization by physiotherapist. Instruction to continue self-treatment sessions for 2 months: (ID, SE, LD, MoE, OP, WCn, BTn) B: Strengthening training: sessions lasting 60–90 min, twice a week for 8 weeks (maximum 15 sessions). Intensive training of trunk flexors and extensors in groups of 6 patients after warm-up exercises. Instruction to continue self-treatment sessions for 2 months: (SD, SE, LD, SgE, OP, WCn, BTn)	Percentage of patients on sick leave after 2, 8 and 14 months FU	No difference after 2 months (A: 10%; B: 14% and 8 months (A: 7%; B: 8%) and 14 months
Roche et al., 2007 (France) (45)	132 patients with chronic LBP on sick leave or at risk of work disability, age 18–50 years	A: Functional restoration programme: 6 h per day, 5 days/week during 5 weeks, muscular warm-up and stretching techniques, aerobic exercises, progressively increasing strengthening exercises, and work simulation. Balneology for 30 min per day. Psychological counselling if required and dietary advice: (ID, SE, HD, MxE, nc, Wcy, BTn) B: Active individual therapy for 1 hour, 3 times a week during 5 weeks and individual exercises to be performed at home for 50 min twice a week: (ID, SE, HD, MxE, OP, WCn, BTn)	RTW and self-assessed ability to work at end of treatment	No difference regarding RTW (A: 85.7%; B: 86.8%). Significant advantage in self-assessed ability to work of A vs B (A: 78.1%; B: 95.5%).

Study	Population	Interventions	Outcome measure	Results
Skouen et al., 2002 (Norway) (46)	195 employees with NSLPB sick-listed > 8 weeks	A: Extensive multidisciplinary programme: 6 h/day, 5 days/week, 4 weeks. Cognitive behavioural modification in group sessions, education, exercise, occasional workplace interventions: (ID, SE, HD, MxE, OP, WCn, BTy) B: Light multidisciplinary programme: information about exercise, fear avoidance. Instruction of a personal exercise programme based on physical tests. FU visits at 3 and 6 months. Further physiotherapy, appointment with psychologist and workplace visits if necessary: (ID, HE, LD, MxE, OP, WCn, BTy) C: Usual care by GP	Sick days (mean, SD) and percentage of patients with full RTW at 12 and 28 months FU	Sign. differences in men for RTW in B vs C but not for women. Overall results: sick days during the first year: A: 167.3; 228.1; B: 149, 228.1; C: 203.8, 223.6); RTW: A: 60%, B: 69%, C: 49%.
Staal et al., 2004, 2005, 2007 (Netherlands) (47–49)	134 sick-listed workers with NSLPB > 4 weeks in succession	A: Graded activity programme: physical examination followed by 1 hour exercise sessions twice per week (graded activity exercises regardless of the amount of pain, work simulation) until complete return to regular work, maximum therapy duration of 3 months: (ID, SE, LD, MxE, OP, WCy, BTy) B: Usual care provided by OP and GP according to the low back guidelines of the Dutch College of General Practitioners	Sick days (mean, SD personal communication) until full return to regular work and subjects sick-listed at 6, 12 and 36 months FU	No significant difference of sick days (mean, SD) at 12mth FU (A: 86.4/90.6; B: 128.5/122.7) and of subjects sick-listed (A: 5; B: 8).
Steenstra et al., 2006 (Netherlands) (50, 51)	112 patients with NSLPB included in a previous trial (44) but remaining sick-listed for > 8 weeks, age 18–65	A: Graded activity programme: individual, submaximal, gradually increasing exercise programme based on the demands from the patients' work with an operant-conditioning behavioural approach consisting of two 1-hour sessions a week, 26 sessions maximally: (ID, SE, HD, MxE, OP, WCy, BTy) B: Usual care provided by OP according to the Dutch OP guidelines for LBP. Patients GP were informed to minimize co-interventions	Persons at work at 26 weeks FU and sick days (mean, SD separated according to previous intervention WI/UC) until full return to regular work at 52 weeks FU	A prolonged RTW at 26 weeks FU (A: 59%; B: 76%), more sick days at 52 weeks (WI: A 160.8, 78.7; B: 109.9, 62.6; UC: A: 172.75, 85.8; B: 151.4, 105.1).
Storheim et al., 2003 (Norway) (52)	93 patients with sub-acute NSLPB, sick-listed for 8–12 weeks	A: Cognitive intervention: 2 consultations, each lasting between 30 and 60 min providing functional examination with individual feedback and advice, explanation of pain mechanisms including reassurance that it is safe to move, and instruction in trunk stabilizing behaviour and squat lifting technique B: Intensive group training: based on the Norwegian Aerobic Fitness Model but modified for patients with LBP lasting 1 hour for 15 weeks, 2–3 times per week. No pain focus: (SD, SE, HD, MxE, OP, WCn, BTy) C: Usual care by the GP	Sick days (mean, SD estimated as 2/3 of mean) at 18 weeks FU	No difference between B and C (B: 63.5, 44; C: 56.3, 37).
Torstensen et al., 1998 (Norway) (53)	137 patients with NSLPB, sick-listed 4–8 weeks, employment, age 20–65 years, birth in Norway	A: Medical exercise therapy: 36 treatments of 1 hour duration for 12 weeks, groups of 5 patients, 6–9 exercises using specially designed exercise equipment with approximately 1000 repetitions: (ID, SE, HD, MxE, OP, WCn, BTn) B: Conventional physiotherapy: 36 treatments of 1 hour duration for 12 weeks, heat or cold, massage, traction, electrotherapy, few exercises on the treatment table: (ID, SE, LD, MxE, OP, WCn, BTn) C: Physician examination and information on the importance of activity and walking 3 times a week	Total number of sick days in each group at 15 months FU	No significant difference regarding sick days between the 3 groups after 15 months (A: 11757, B: 9967, C: 13567).
White, 1966 (Canada) (54)	194 men sick-listed 6–52 weeks for NSLPB, age 19–60 years	A: Inpatient rehabilitation: up to 6 weeks, 7 h per day, progressive treatment in 4 stages (bed rest–light–medium–heavy physical exercise) to the level demanded by the requirements of the man's job: (ID, SE, HD, MxE, IP, WCy, BTn) B: Usual care by GP	Numbers of workers with satisfactory RTW at 3 months FU	Significantly better at 3 months (A: 42%, B: 16%.

ID: individually designed; SD: standard design; HE: home exercises; SE: supervised exercise; HD: high-dose exercise; LD: low-dose exercise; MxE: mixed exercises; SgE: strengthening exercises; ScE: stretching exercises; SaE: stabilization exercises; MoE: mobilization exercises; IP: inpatient; OP: outpatient; WCy/n: work context yes/no; BTy/n: behavioural treatment yes/no; FU: follow-up; LBP: low back pain; RTW: return to work.

RESULTS

Trial flow

Of the 838 articles retrieved from the literature search, we evaluated 87 articles in detail, of which 64 did not meet the inclusion criteria. Consequently, we included 23 studies in this review. Sixteen were trials with 2 study arms and 7 were trials with 3 study arms. Table II summarizes the characteristics of the included studies. Twenty studies were included in the meta-analysis, allowing 17 comparisons of exercise interventions with usual care (5, 27, 29, 31, 32, 35, 41, 46, 47, 50, 52–54) and 11 comparisons of 2 different exercise interventions (20, 21, 29, 31, 33, 34, 37, 43, 45, 46, 53). Three studies (22, 25, 26) were excluded from meta-analysis as "days of sick leave" were presented as median and interquartile range, thereby preventing pooling (Fig. 1).

Trials comparing 2 different exercise interventions with usual care were treated as 2 trials with the sample size of the usual care group equally divided between the 2 exercise intervention groups: inpatient rehabilitation and outpatient treatment (29); low and high intensity back school (31); light and extensive multidisciplinary programme (46); conventional physiotherapy and medical exercise therapy (53). One study (50) presented results of 2 patient groups defined by the previous intervention (UC: usual care; WI: workplace intervention) receiving the same exercise intervention.

Validity assessment

According to the previously mentioned criteria, 14 (61%) of the studies were found to be of high quality and 9 (39%) of low
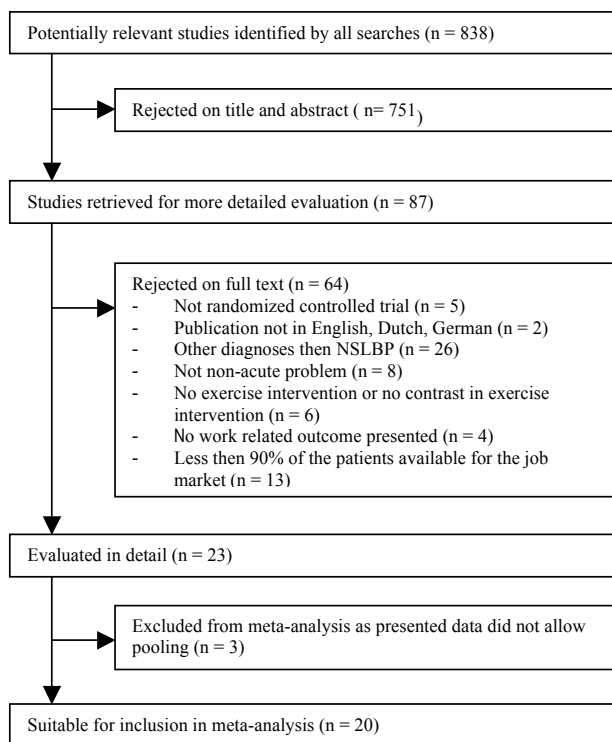
quality (Table III). The 3 studies excluded from meta-analysis were all of low quality.

Work disability data

Data on work disability varied between the different studies and included self-assessed work ability, days of sick leave, days at work, physician's judgement of work capability, and days of sickness compensation or numbers of workers returning to full-duty work. These were obtained from insurance databases whereby national legal requirements may have influenced the recordings. The data used for pooling were the number of people who returned and did not returned to work at the time of the follow-up, or the total number of sick days within the follow-up period (Table II).

Exercise characteristics

Thirty-five different exercise interventions were used. Exercise design, dose and setting were reported unclear in 6% of the investigated exercise interventions. Twenty-six (74%) of the exercise interventions were individually designed; 32 (91%) were primarily performed as supervised exercise; 28 (80%) interventions used mixed exercise types, 2 stabilization, 3 strengthening, 1 mobilization, and 1 stretching exercise; 27 (77%) were conducted in an outpatient setting; 10 (29%) were work-related; and 14 (40%) of the exercise interventions were administered within a cognitive behavioural approach (Table II).

Unfortunately, none of the studies using home exercise reported adherence rate or sufficient information to estimate home exercise dose. Therefore, the calculation of the exercise dose is based on the number of the supervised treatment ses-

Fig. 1. Trial flow diagram to summarize the stages of the systematic review.

Table III. *Risk of bias and study quality*

Study, reference	Risk of bias			Study quality
	Selection bias	Detection bias	Attrition bias	
Alaranta et al., 1994 (20)	Yes	No	Yes	Low
Aure et al., 2003 (21)	No	Yes	No	High
Bendix et al.,1995 (22–24)	Yes	Yes	Yes	Low
Bendix et al.,1996 (24, 25)	Yes	Yes	Yes	Low
Bendix et al., 2000 (26)	Yes	Yes	Yes	Low
Hagen et al., 2000 (27, 28)	No	No	No	High
Härkäpää et al., 1989 (29, 30)	Yes	No	Yes	Low
Heymans et al., 2006 (31)	No	No	No	High
Hurri, 1989 (32)	Yes	Yes	No	Low
Jousset et al., 2004 (32)	Yes	Yes	Yes	Low
Kääpa et al., 2006 (34)	No	Yes	No	High
Karjalainen et al.,2003 (35, 36)	No	Yes	No	High
Kool et al., 2005 (37, 38)	No	No	No	High
Lindström et al., 1992 (5, 39, 40)	No	No	No	High
Niemisto et al., 2003 (41, 42)	No	Yes	No	High
Petersen et al. 2002 (43, 44)	No	Yes	No	High
Roche et al., 2007 (45)	Yes	Yes	No	Low
Skouen et al., 2002 (46)	No	No	Yes	High
Staal et al., 2004 (47–49)	No	No	No	High
Steenstra et al., 2006 (50, 51)	No	No	No	High
Storheim et al., 2003 (52)	No	No	No	High
Torstensen et al., 1998 (53)	No	No	No	High
White, 1966 (54)	Yes	Yes	Yes	Low

sions and their duration only. Such calculated exercise dose varied widely between the different exercise interventions, ranging between 1.5 and 210 h. The median exercise dose was 17 h. We classified exercise interventions with ≥ 17 h of contact time as high-dose exercise ($n = 18$), and those with less than 17 h of contact time into low-dose exercises ($n = 17$). A cut-off point of 14 and 20 h resulted in less than a 10% change in exercise dose classification.

Qualitative comparison of exercise interventions and usual care

Seven studies reported work disability data on a short-term follow-up. Two high-quality studies (5, 27) and 2 low-quality studies (25, 54) reported a positive effect, 1 high-quality study no effect (52) and 2 high-quality studies a negative effect (31, 50). Five studies reported work disability data on an intermediate-term follow-up. Three high-quality studies (5, 27, 47) reported a positive effect. This was observed in the study by Staal et al. (47) from approximately 50 days after randomization onwards. Two high-quality studies reported a negative effect (31, 50). Long-term results were presented by 11 studies. Positive effects were found in 3 high-quality studies (5, 27, 35). No significant effects were observed in 3 low-quality studies

(25, 29, 32) and 4 high-quality studies (41, 46, 47, 53). One high-quality study reported a negative effect (50).

Quantitative data synthesis

Comparison of exercise interventions vs usual care. Thirteen studies allowing 17 comparisons between an exercise intervention and usual care with a total of 3181 patients were available for pooling.

Comparison 01: Short-term follow-up. Short-term results were available for pooling from 5 high-quality studies (6 comparisons, 1030 patients) (5, 27, 31, 50, 52), showing no significant effect of exercise reducing work disability (OR = 0.80, 95% confidence interval (CI) 0.51–1.25). The addition of one low-quality study (54) did not substantially change the overall effect estimate (OR = 0.68, 95% CI 0.42–1.10)

Comparison 02: Intermediate-term follow-up. Four high-quality studies (5 comparisons, 971 patients) (5, 27, 31, 50) provided results for pooling at the intermediate-term follow-up showing no significant effect of exercise in reducing work disability (OR = 0.78, 95% CI 0.45–1.34).
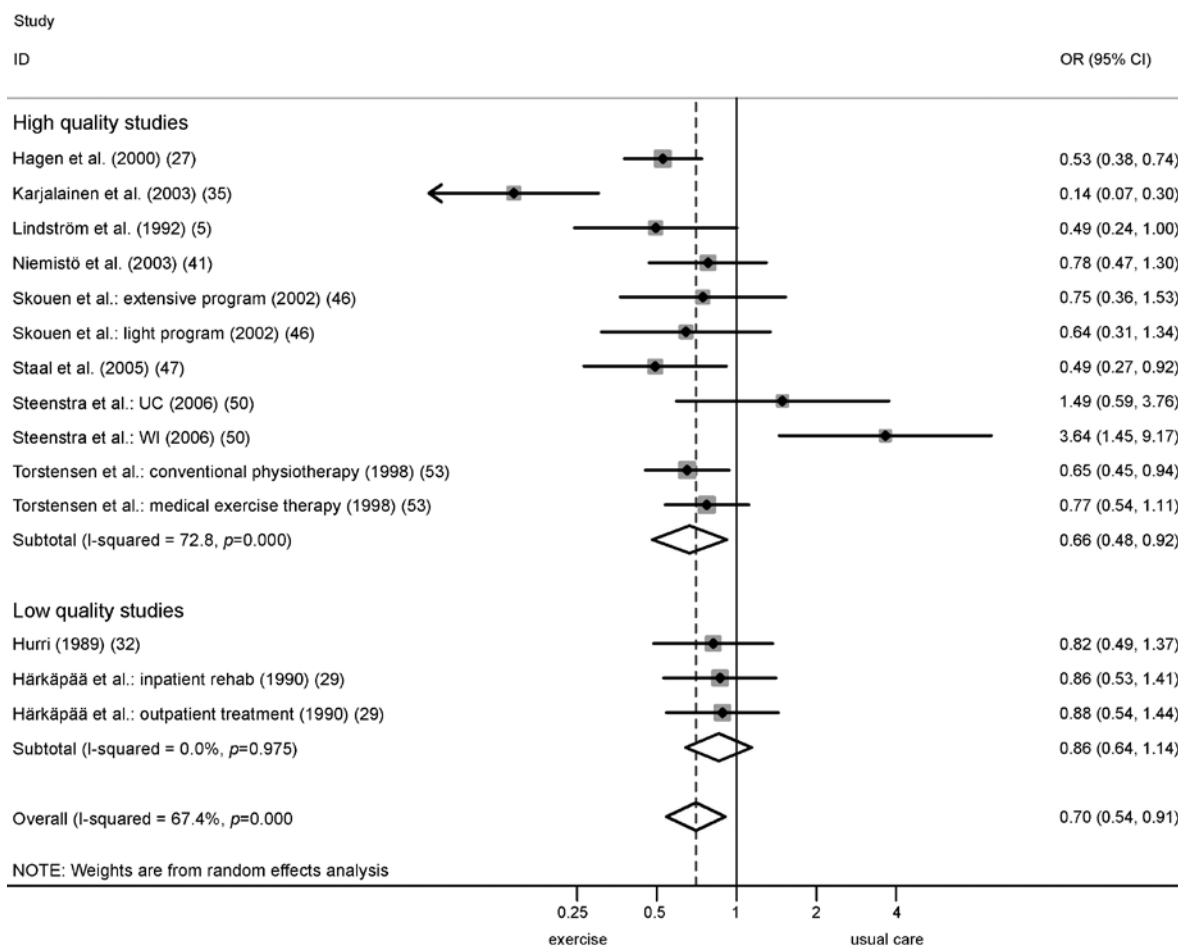

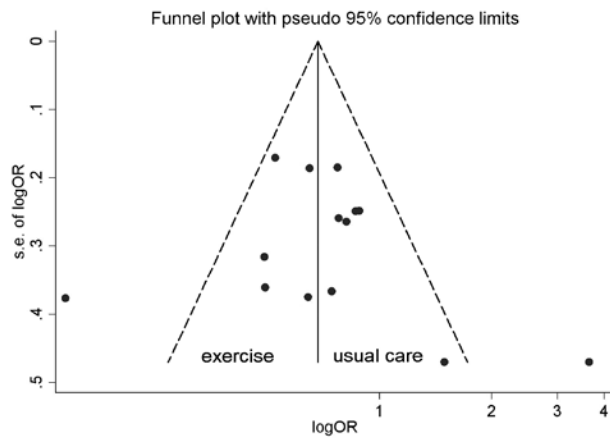*Fig. 2.* Meta-analysis of 10 trials with long-term follow-up comparing exercise interventions with usual care.

Funnel plot with pseudo 95% confidence limits

Fig. 3. Funnel plot of 10 trials with long-term follow-up comparing exercise interventions with usual care.

Comparison 03: Long-term follow-up. Eight high-quality studies (10 comparisons, 1992 patients (5, 27, 35, 41, 46, 47, 50, 53) presented long-term follow-up results showing a statistically significant overall effect in favour of exercise on work disability (OR = 0.66, 95% CI 0.48–0.92). The addition of 2 low-quality studies (29, 32) did not substantially change the overall effect estimate (OR = 0.70, 95% CI 0.54–0.91) (Fig. 2). The funnel plots did reveal evidence of asymmetry at short- and intermediate-term follow-ups but not at long-term follow-up (Fig. 3).

Comparison 04: Influence of exercise characteristics in high-quality trials with long-term follow-up. The 8 high-quality studies (5, 27, 35, 41, 46, 47, 50, 53) providing data on 1149 patients receiving an exercise intervention and 843 patients receiving usual care were included for this analysis. All comparisons were between different outpatient rehabilitation programmes, and all used individually designed exercises, in one comparison stretching exercises were instructed, another used stabilization exercise, and in the remaining 8 comparisons

mixed exercises were used. A second overall analysis, which did not include the patient sample from the trial of Steenstra et al. (50) that had already received a workplace intervention (WI group), showed reduced statistical heterogeneity ($I^2 = 60.4\%$, $p = 0.007$) and increased the effect estimate (OR = 0.59, 95% CI 0.45–0.78). The effect of delivery type, exercise dose, work context and behavioural treatment approach was analysed with and without the WI group showing different results, although within the statistical error margin. Pooled effects for the 4 exercise characteristics hypothesized to influence work disability (delivery type, exercise dose, work context, behavioural treatment approach) became higher and more significant for supervised exercise, and a behavioural treatment approach. However, none of the variables were statistically significant in meta-regression analysis, although there was a trend observed favouring home exercises ($p = 0.11$) (Table IV).

Comparison of different exercise interventions. Of the 13 studies comparing 15 different exercise interventions, 6 were of low quality (20, 22, 26, 29, 33, 45) and 7 were of high quality. Six high-quality studies presented long-term follow-up data and were used for pooling (21, 34, 37, 43 46, 53). We defined the exercise intervention with more contact hours as the standard intervention. All standard interventions used individually designed supervised exercises, 5 of them with mixed exercises, and 5 were conducted in an outpatient setting.

There was significant statistical heterogeneity in these trials (I-squared = 65.5%, $p = 0.013$). The overall effect of exercise interventions with more contact hours was not significant (OR = 1.07, 95% CI 0.67–1.72). Three trials applying exercise within a behavioural treatment approach showed some benefit (OR = 0.75, 95% CI 0.47–1.20) compared with the trials without this characteristic (OR = 1.74, 95% CI 0.71–4.30) (Fig. 4). One trial applying work-related exercise in an inpatient setting (42) showed a significant effect on work disability (OR = 0.53, 95% CI 0.30–0.93) compared with exercise not specifically designed to restore work-related physical capacity (OR = 1.25, 95% CI 0.80–1.97). None of these characteristics showed statistical significance in meta-regression analysis.

Table IV. *Odds ratios for work disability stratified by exercise characteristics in 8 high-quality randomized controlled trials comparing exercise intervention with usual care*

Exercise characteristics	Work disability (WI group included) OR (95% CI)	Work disability (WI group excluded) OR (95% CI)	Meta-regression (WI group excluded) logOR (95% CI), *p*-value
Delivery type			
Home exercises	0.38 (0.17–0.84)	0.38 (0.17–0.84)	1.74 (0.86–3.55), *p* = 0.11
Supervised exercise	0.80 (0.58–1.11)	0.70 (0.58–0.85)	
Exercise dose			
Low dose (<17 hours)	0.51 (0.35–0.73)	0.51 (0.35–0.73)	1.52 (0.71–3.27), *p* = 0.24
High dose (>17 hours)	1.01 (0.57–1.78)	0.76 (0.56–1.05)	
Work context			
No	0.65 (0.54–0.77)	0.65 (0.54–0.77)	0.66 (0.27–1.59), *p* = 0.31
Yes	0.77 (0.21–2.85)	0.46 (0.41–1.55)	
Behavioural treatment approach			
No	0.72 (0.57–0.91)	0.72 (0.57–0.91)	0.72 (0.33–1.56), *p* = 0.35
Yes	0.65 (0.39–1.10)	0.52 (0.34–0.80)	

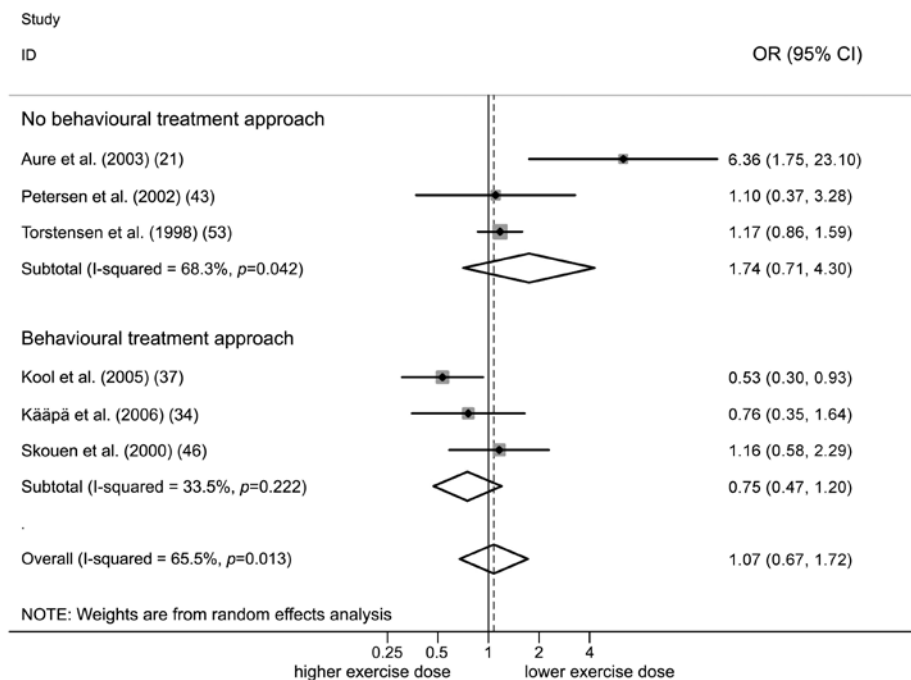CI: confidence interval; OR: odds ratio; WI: workplace intervention.

Fig. 4. Meta-analysis of 6 high-quality trials comparing lower and higher dosed exercises at long-term follow-up without and with a behavioural treatment approach.

The funnel plots that were conducted did not reveal evidence of funnel plot asymmetry.

DISCUSSION

This meta-analysis provides continuous support for the use of exercise interventions to achieve long-term benefits on work disability in patients with non-acute NSLBP. The OR of 0.66 suggests that the odds of "improvement" in work disability are in the long-term 34% lower if only usual care (rather than exercise) is given. No significant effect was observed in short- and intermediate-term follow-ups. Meta-regression showed no significant differences between different exercise types. Interestingly, home exercises seem to be at least as effective as supervised programmes. As the meta-regression is only explorative, no conclusions can be made regarding exercise types.

Our study has several strengths. The search strategy was based on the recommendations of the Cochrane Back Review Group. We planned the analysis *a priori* based on the findings of previous meta-analyses and assessed study quality based on key components of methodological quality (concealed allocation, blinded assessor, intention to treat analysis) as recommended by Juni et al. (13). Studies affected by biases have previously been shown to exaggerate treatment effects (13). We, therefore, excluded low-quality studies from meta-regression analysis in order to avoid a possible overestimation of the effect of different exercise characteristics.

A weakness of this study is the high proportion of total unexplained variance that could be attributed to study heterogeneity.

We considered this weakness thoroughly, but concluded that patients, social support and outcomes showed satisfactory homogeneity. All but one of the studies were performed in Europe, in countries with comparable social systems. All patients were diagnosed with non-acute NSLBP, were of working age and available for the job market. Despite the wide variety of used work disability outcomes, this meta-analysis is based on the pooled results of just 2 different outcome measures. We performed a stratified analysis in 3 studies providing both outcome measures and found no relevant differences in ORs, both in favour of exercise. Furthermore, using mean values and standard deviations for further statistical analysis in data with a skewed distribution is usually regarded as inappropriate. Data regarding sick days have a skewed distribution, but this was similar in both groups in treatment comparisons that reduces the risk of systematic bias (55). To address the problem of statistical heterogeneity, we performed a random effects meta-analysis. There remains the possible error of substantial variation in standard deviations across studies leading to an over- or underestimation of the ORs.

Exercise interventions did not show a significant effect on work disability at short- and intermediate-term follow-up. However, these findings are not conclusive. The mean odds ratios for short- and intermediate-term results were both below unity, but with wide CI. Therefore, a significant effect might remain undetected based on ineffectiveness, heterogeneity or limited power of the pooled studies. Possible explanations for a lack of effect at short- and intermediate-term follow-up are the required time needed to improve physical capacity, to modify pain behaviour, or to search for work. Furthermore,

the process of care has a substantial effect on work disability, as shown in a recent study comparing a graded activity programme with usual care (50). The interaction between a prior WI and graded activity, together with a delay in the start of the graded activity intervention, explained most of the delay in return to work (RTW) (50). This study introduced relevant clinical heterogeneity in this meta-analysis. All of the other trials investigated the primary treatments for this occurrence of back pain, while half of the patients in the trial from Steenstra et al. (50) had already received a WI, which has been shown to be effective on return to work (56). Herbert & Bo (57) propose that study quality can also be assessed on how interventions are administered. There were obviously problems in the implementation of the graded activity programme in the trial of Steenstra et al. (50), leading to a potentially false conclusion if the whole patient sample had been included in this meta-analysis. In view of these considerations we feel it legitimate to interpret the findings without the results of the WI group. The author's recommendation of paying special attention to the structure and process of care in implementing graded activity (50) does have clinical relevance when conducting medical interventions aiming for early RTW. An open and fast access to such interventions prevents unnecessary waiting time before a RTW can be attempted. This might also be a possible explanation for why individually designed home exercises seem to be more effective than supervised exercise interventions in reducing work disability. Home exercise may facilitate RTW, as the patients are able to continue their daily routine without spending extra time on medical intervention.

Interestingly, this meta-analysis did not show a greater effect of higher dose exercise interventions (≥ 17 contact h) compared with lower dose exercise interventions (< 17 contact h) on work disability. This finding is contrary to exercise physiology postulating a dose and effect relation (58), as well as to previous findings that only intensive (> 100 h of therapy) multidisciplinary biopsychosocial rehabilitation with functional restoration improves function in patients with chronic low back pain, whereby inconclusive results were found on vocational outcomes (59). The only study included in the systematic review by Guzman et al. (59) supporting the use of functional restoration to reduce work disability (25) is not included in the performed meta-regression, as only high-quality studies were used. Moreover, in the meantime new studies with low contact hours administering home exercise have been published showing a positive effect on work disability. This might, from our point of view, be an explanation for the different findings. However, a cautionary comment must be made on the missing effect of exercise dose found in this review. As in other systematic reviews, incomplete reporting in the primary studies present important limitations and prevented the calculation of the exercise dose in home exercise programmes. It must be assumed that the actual exercise dose in home exercise interventions was higher than the calculated dose.

We have not been able to confirm the positive effects of exercises performed within a behavioural treatment approach on work disability postulated in previous reviews (8, 9, 14). In the comparison of exercises with usual care we found stronger

treatment effects for such exercises. However, this was not statistically significant in the meta-regression (Table IV). The missing confirmation might be due to the differing study inclusion criteria and the analysis performed. All previous reviews based their conclusion on a qualitative assessment, at least partly based on the evidence found by Lindström et al. (5). Hayden et al. (9) also included the results of Staal et al. (47), while Schonstein et al. 2003 (14) included the results of 3 more studies that were excluded from this analysis because of the risk of bias or missing inclusion criteria. We included in addition to the studies of Staal et al. (47) and Lindström et al. (5) the findings of 6 more studies (27, 31, 35, 47, 50, 52) that contained 9 treatment comparisons with a total of 1316 patients providing sufficient power for the meta-regression. However, it must be emphasized that the presented meta-regression analyses is only explorative and does not allow any conclusions too be drawn.

The comparison of different exercise interventions also did not reveal a significant effect of a behavioural treatment approach (Fig. 4). There might be a superior effect if exercises are performed within a behavioural treatment approach and are specifically designed to restore work-related capacity, as shown by the study of Kool et al. (38). This is also in line with Schonstein et al. (7), who hypothesize a positive effect of such a combination.

We recommend further evaluation of the combined effects of individually designed home exercises applied within a behavioural treatment approach aiming to specifically restore work-related physical capacity. Special attention must be given to an effective implementation process of exercise interventions aiming for early RTW.

ACKNOWLEDGEMENTS

REFERENCES

1. Woolf AD, Pfleger B. Burden of major musculoskeletal conditions. Bull World Health Organ 2003; 81: 646–656.
2. Abenhaim L, Rossignol M, Valat JP, Nordin M, Avouac B, Blotman F, et al. The role of activity in the therapeutic management of back pain. Report of the International Paris Task Force on Back Pain. Spine 2000; 25 Suppl 4: 1S–33S.
3. COST B13. European guidelines for the management of acute nonspecific low back pain in primary care. Eur Spine J 2006; 15 Suppl 2: S125–S300.
4. Mayer T, Gatchel R. Functional Restoration for Spinal Disorders: The Sports Medicine Approach. Philadelphia: Lea and Febiger; 1988.
5. Lindström I, Öhlund C, Eek C, Wallin L, Peterson LE, Nachemson A. Mobility, strength, and fitness after a graded activity program for patients with subacute low back pain. A randomized prospective clinical study with a behavioral therapy approach. Spine 1992; 17: 641–652.
6. Hayden JA, van Tulder MW, Tomlinson G. Systematic review: strategies for using exercise therapy to improve outcomes in

chronic low back pain. Ann Intern Med 2005; 142: 776–785.

7. Schonstein E, Kenny DT, Keating J, Koes BW. Work conditioning, work hardening and functional restoration for workers with back and neck pain. Cochrane Database Syst Rev 2003; (1): CD001822.

8. Ostelo R, van Tulder M, Vlaeyen J, Linton S, Morley S, Assendelft W. Behavioural treatment for chronic low-back pain. Cochrane Database Syst Rev 2005: CD002014. DOI: 10.1002/14651858. CD002014.pub2(1).

9. Hayden JA, van Tulder MW, Malmivaara AV, Koes BW. Meta-analysis: exercise therapy for nonspecific low back pain. Ann Intern Med 2005; 142: 765–775.

10. Kool J, de Bie R, Oesch P, Knusel O, van den Brandt P, Bachmann S. Exercise reduces sick leave in patients with non-acute non-specific low back pain: a meta-analysis. J Rehabil Med 2004; 36: 49–62.

11. Hlobil H, Staal JB, Spoelstra M, Ariens GA, Smid T, van Mechelen W. Effectiveness of a return-to-work intervention for subacute low-back pain. Scand J Work Environ Health 2005; 31: 249–257.

12. van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. Spine 2003; 28: 1290–1299.

13. Juni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials. BMJ 2001; 323: 42–46.

14. Schonstein E, Kenny D, Keating J, Koes B, Herbert RD. Physical conditioning programs for workers with back and neck pain: a Cochrane systematic review. Spine 2003; 28: E391–E395.

15. Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. Stat Med 2000; 19: 3127–3131.

16. Hasselblad V, Hedges LV. Meta-analysis of screening and diagnostic tests. Psychol Bull 1995; 117: 167–178.

17. Der Simonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986; 7: 177–188.

18. Reitman RD, Emerson RH, Higgins LL, Tarbox TR. A multimodality regimen for deep venous thrombosis prophylaxis in total knee arthroplasty. J Arthroplasty 2003; 18: 161–168.

19. Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. J Clin Epidemiol 2001; 54: 1046–1055.

20. Alaranta H, Rytokoski U, Rissanen A, Talo S, Ronnemaa T, Puukka P, et al. Intensive physical and psychosocial training program for patients with chronic low back pain. A controlled clinical trial. Spine 1994; 19: 1339–1349.

21. Aure OF, Nilsen JH, Vasseljen O. Manual therapy and exercise therapy in patients with chronic low back pain: a randomized, controlled trial with 1-year follow-up. Spine 2003; 28: 525–531; discussion 31–32.

22. Bendix AF, Bendix T, Ostenfeld S, Bush E, Andersen. Active treatment programs for patients with chronic low back pain: a prospective, randomized, observer-blinded study. Eur Spine J 1995; 4: 148–152.

23. Bendix AF, Bendix T, Lund C, Kirkbak S, Ostenfeld S. Comparison of three intensive programs for chronic low back pain patients: a prospective, randomized, observer-blinded study with one-year follow-up. Scand J Rehabil Med 1997; 29: 81–89.Hagen EM, Eriksen HR, Ursin H. Does early intervention with a light mobilization program reduce long-term sick leave for low back pain? Spine 2000; 25: 1973–1976.

24. Bendix AF, Bendix T, Labriola M, Boekgaard P. Functional restoration for chronic low back pain. Two-year follow-up of two randomized clinical trials. Spine 1998; 23: 717–725.

25. Bendix AF, Bendix T, Vaegter K, Lund C, Frolund L, Holm L. Multidisciplinary intensive treatment for chronic low back pain: a randomized, prospective study. Cleve Clin J Med 1996; 63: 62–69.

26. Bendix T, Bendix A, Labriola M, Haestrup C, Ebbehoj N. Functional restoration versus outpatient physical training in chronic low back pain: a randomized comparative study. Spine 2000; 25: 2494–2500.

27. Hagen EM, Eriksen HR, Ursin H. Does early intervention with a light mobilization program reduce long-term sick leave for low back pain? Spine 2000; 25: 1973–1976.

28. Hagen EM, Grasdal A, Eriksen HR. Does early intervention with a light mobilization program reduce long-term sick leave for low back pain: a 3-year follow-up study. Spine 2003; 28: 2309–2315.

29. Härkäpää K, Jarvikoski A, Mellin G, Hurri H. A controlled study on the outcome of inpatient and outpatient treatment of low back pain. Part I. Pain, disability, compliance, and reported treatment benefits three months after treatment. Scand J Rehabil Med 1989; 21: 81–89.

30. Härkäpää K, Mellin G, Jarvikoski A, Hurri H. A controlled study on the outcome of inpatient and outpatient treatment of low back pain. Part III. Long-term follow-up of pain, disability, and compliance. Scand J Rehabil Med 1990; 22: 181–188.

31. Heymans MW, de Vet HC, Bongers PM, Knol DL, Koes BW, van Mechelen W. The effectiveness of high-intensity versus low-intensity back schools in an occupational setting: a pragmatic randomized controlled trial. Spine 2006; 31: 1075–1082.

32. Hurri H. The Swedish back school in chronic low back pain. Part I. Benefits. Scand J Rehabil Med 1989; 21: 33–40.

33. Jousset N, Fanello S, Bontoux L, Dubus V, Billabert C, Vielle B, et al. Effects of functional restoration versus 3 hours per week physical therapy: a randomized controlled study. Spine 2004; 29: 487–493; discussion 494.

34. Kääpä EH, Frantsi K, Sarna S, Malmivaara A. Multidisciplinary group rehabilitation versus individual physiotherapy for chronic nonspecific low back pain: a randomized trial. Spine 2006; 31: 371–376.

35. Karjalainen K, Malmivaara A, Pohjolainen T, Hurri H, Mutanen P, Rissanen P, et al. Mini-intervention for subacute low back pain: a randomized controlled trial. Spine 2003; 28: 533–540; discussion 40–41.

36. Karjalainen K, Malmivaara A, Mutanen P, Roine R, Hurri H, Pohjolainen T. Mini-intervention for subacute low back pain: two-year follow-up and modifiers of effectiveness. Spine 2004; 29: 1069–1076.

37. Kool JP, Oesch PR, Bachmann S, Knuesel O, Dierkes JG, Russo M, et al. Increasing days at work using function-centered rehabilitation in nonacute nonspecific low back pain: a randomized controlled trial. Arch Phys Med Rehabil 2005; 86: 857–864.

38. Kool J, Bachmann S, Oesch P, Knuesel O, Ambergen T, de Bie R, et al. Function-centered rehabilitation increases work days in patients with nonacute nonspecific low back pain: 1-year results from a randomized controlled trial. Arch Phys Med Rehabil 2007; 88: 1089– 1094.

39. Lindström I, Ohlund C, Nachemson A. Physical performance, pain, pain behavior and subjective disability in patients with subacute low back pain. Scand J Rehabil Med 1995; 27: 153–160.

40. Lindström I, Öhlund C, Eek C, Wallin L, Peterson LE, Fordyce WE, et al. The effect of graded activity on patients with subacute low back pain: a randomized prospective clinical study with an operant-conditioning behavioral approach. Phys Ther 1992; 72: 279–290.

41. Niemisto L, Lahtinen-Suopanki T, Rissanen P, Lindgren KA, Sarna S, Hurri H. A randomized trial of combined manipulation, stabilizing exercises, and physician consultation compared to physician consultation alone for chronic low back pain. Spine 2003; 28: 2185–2191.

42. Niemisto L, Rissanen P, Sarna S, Lahtinen-Suopanki T, Lindgren KA, Hurri H. Cost-effectiveness of combined manipulation, stabilizing exercises, and physician consultation compared to physician consultation alone for chronic low back pain: a prospective randomized trial with 2-year follow-up. Spine 2005; 30: 1109–1115.

43. Petersen T, Kryger P, Ekdahl C, Olsen S, Jacobsen S. The effect of McKenzie therapy as compared with that of intensive strengthening training for the treatment of patients with subacute or chronic

low back pain: a randomized controlled trial. Spine 2002; 27: 1702–1709.

44. Petersen T, Larsen K, Jacobsen S. One-year follow-up comparison of the effectiveness of McKenzie treatment and strengthening training for patients with chronic low back pain: outcome and prognostic factors. Spine 2007; 32: 2948–2956.

45. Roche G, Ponthieux A, Parot-Shinkel E, Jousset N, Bontoux L, Dubus V, et al. Comparison of a functional restoration program with active individual physical therapy for patients with chronic low back pain: a randomized controlled trial. Arch Phys Med Rehabil 2007; 88: 1229–1235.

46. Skouen JS, Grasdal AL, Haldorsen EM, Ursin H. Relative cost-effectiveness of extensive and light multidisciplinary treatment programs versus treatment as usual for patients with chronic low back pain on long-term sick leave: randomized controlled study. Spine 2002; 27: 901–909.

47. Staal JB, Hlobil H, Twisk JW, Smid T, Koke AJ, van Mechelen W. Graded activity for low back pain in occupational health care: a randomized, controlled trial. Ann Intern Med 2004; 140: 77–84.

48. Hlobil H, Staal JB, Twisk J, Köke A, Ariëns G, Smid T, et al. The effects of a graded activity intervention for low back pain in occupational health on sick leave, functional status and pain: 12-month results of a randomized controlled trial. J Occup Rehabil 2005; 15: 569–580.

49. Hlobil H, Uegaki K, Staal JB, de Bruyne MC, Smid T, van Mechelen W. Substantial sick-leave costs savings due to a graded activity intervention for workers with non-specific sub-acute low back pain. Eur Spine J 2007; 16: 919–924.

50. Steenstra IA, Anema JR, Bongers PM, de Vet HC, Knol DL, van Mechelen W. The effectiveness of graded activity for low back pain in occupational healthcare. Occup Environ Med 2006; 63: 718–725.

51. Steenstra IA, Anema JR, van Tulder MW, Bongers PM, de Vet HC, van Mechelen W. Economic evaluation of a multi-stage return to work program for workers on sick-leave due to low back pain. J Occup Rehabil 2006; 16: 557–578.

52. Storheim K, Brox JI, Holm I, Koller AK, Bo K. Intensive group training versus cognitive intervention in sub-acute low back pain: short-term results of a single-blind randomized controlled trial. J Rehabil Med 2003; 35: 132–140.

53. Torstensen TA, Ljunggren AE, Meen HD, Odland E, Mowinckel P, Geijerstam S. Efficiency and costs of medical exercise therapy, conventional physiotherapy, and self-exercise in patients with chronic low back pain. A pragmatic, randomized, single-blinded, controlled trial with 1-year follow-up. Spine 1998; 23: 2616–2624.

54. White AW. Low back pain in men receiving workmen's compensation. Can Med Assoc J 1966; 95: 50–56.

55. Boneau C. The effects of violations of assumptions underlying the t-test. Psychol Bull 1960; 57: 49–64.

56. Anema JR, Steenstra IA, Bongers PM, de Vet HC, Knol DL, Loisel P, et al. Multidisciplinary rehabilitation for subacute low back pain: graded activity or workplace intervention or both: a randomized controlled trial [with consumer summary]. Spine 2007; 32: 291–298.

57. Herbert RD, Bo K. Analysis of quality of interventions in systematic reviews. BMJ 2005; 331: 507–509.

58. Wilmore J, Costill D, Kenney W. Physiology of sport and exercise science, 4th edn. Hong Kong: Human Kinetics; 2008.

59. Guzman J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary rehabilitation for chronic low back pain: a systematic review. BMJ 2001; 2001; 322: 1511–1516.