Pain 107 (2004) 176–190

PAIN

www.elsevier.com/locate/pain

Exercise and chronic low back pain: what works?

Sarah Dianne Liddle*, G. David Baxter, Jacqueline H. Gracey

Rehabilitation Sciences Research Group, University of Ulster at Jordanstown, Shore Road, Newtownabbey, Co. Antrim BT37 OQB, Northern Ireland, UK

Received 22 July 2003; received in revised form 16 October 2003; accepted 20 October 2003

Abstract

The aim of this review was to investigate current evidence for the type and quality of exercise being offered to chronic low back pain (CLBP) patients, within randomised controlled trials (RCTs), and to assess how treatment outcomes are being measured. A two-fold methodological approach was adopted: a methodological assessment identified RCTs of 'medium' or 'high' methodological quality. Exercise quality was subsequently assessed according to the predominant exercise used. Outcome measures were analysed based on current recommendations. Fifty-four relevant RCTs were identified, of which 51 were scored for methodological quality. Sixteen RCTs involving 1730 patients qualified for inclusion in this review based upon their methodological quality, and chronicity of symptoms; exercise had a positive effect in all 16 trials. Twelve out of 16 programmes incorporated strengthening exercise, of which 10 maintained their positive results at follow-up. Supervision and adequate compliance were common aspects of trials. A wide variety of outcome measures were used. Outcome measures did not adequately represent the guidelines for impairment, activity and participation, and impairment measures were over-represented at the expense of others. Despite the variety offered, exercise has a positive effect on CLBP patients, and results are largely maintained at follow-up. Strengthening is a common component of exercise programmes, however, the role of exercise co-interventions must not be overlooked. More high quality trials are needed to accurately assess the role of supervision and follow-up, together with the use of more appropriate outcome measures.

Keywords: Chronic low back pain; Exercise; Outcome measures; Supervision; Compliance; Methodological quality

1. Introduction

During the course of their lives 70–85% of individuals will experience low back pain (LBP) (Andersson, 1999; Deyo and Weinstein, 2001; Goodwin and Goodwin, 2000; van Tulder, 2001); furthermore, over 80% of such patients report recurrent episodes (Waddell, 1998, p. 103–117). It is estimated that 80–90% of patients will have recovered within 6 weeks, regardless of treatment (Bronfort et al., 1996; CSAG, 1994; Indahl et al., 1995; Jackson, 2001; Klaber Moffett et al., 1986; Lahad et al., 1994; van Tulder et al., 1997). However, 5–15% will develop chronic low back pain (CLBP; >12 weeks) (Bigos et al., 2001; Johannsen et al., 1995; Klaber Moffett et al., 1986; Quittan, 2002; Tortensen et al., 1998): this is more difficult to treat (Cottingham and Maitland, 1997; Frost et al., 2000; Hildebrandt et al., 1997) and treatment has variable results (Carpenter and Nelson, 1999; CSAG, 1994; Rainville et al., 1997).

van Tulder et al. (1997) have highlighted that clinical guidelines are needed for the management of CLBP in primary care. Current evidence suggests that exercise and intensive multidisciplinary treatment programmes are likely to be beneficial for CLBP. Exercise is thought to decrease fear-avoidance behaviour and facilitate functional improvements, despite ongoing pain. This is an important component of the popular biopsychosocial model of CLBP management (Frost et al., 2000; Hartigan et al., 2000; Lively, 2002; Mannion et al., 1999; Pfingsten, 2001; Rainville et al., 1997); this notwithstanding, there does not appear to be a consensus of opinion on the most effective programme design to maintain exercise benefits (Bronfort et al., 1996; Carpenter and Nelson, 1999; Faas, 1996; Kenny, 2000; Lahad et al., 1994; Ljunggren et al., 1997; Manniche et al., 1991; Taimela and Harkapaa, 1996; Taimela et al., 2000). Long-term maintenance of these benefits requires patient education and motivation towards behavioural change and exercise

* Corresponding author. Tel.: +7773-851983; fax: +2890-366839.
E-mail address: sd.liddle@ulster.ac.uk (S.D. Liddle).

0304-3959/$20.00
doi:10.1016/j.pain.2003.10.017

compliance (ACSM, 2000, p. 245). Programme supervision is thought to play a part in enhancing exercise compliance (ACSM, 2000 p. 162).

The value of outcome measures for assessing the effectiveness of various therapies has been highlighted in the literature (Beattie and Maher, 1997; Bombardier, 2000; Koes et al., 1995). Bombardier (2000) has identified a core set of outcome measures that includes the following five domains: back specific function, generic health status, pain, work disability and patient satisfaction. The authors acknowledge that there is no 'ideal' set of outcome measures, however, the guidelines proposed by Bombardier (2000) are considered representative of the biopsychosocial influences on the CLBP patient and have therefore been followed throughout this review.

2. Aims of the review

To date, systematic reviews in this area have failed to address the clinical relevance of type, quality, and mode of delivery of exercise. In addition, insufficient attention has been given to promoting the consistent use of well-validated outcome measures. The fundamental objective of this review was to identify what treatment characteristics are essential to achieve and maintain successful results with exercise in this group of pain patients and to investigate how treatment outcomes are being measured.

3. Methodology

3.1. Overview

The QUOROM guidelines (Moher et al., 1999) were used as a methodological template for this review. A two-stage methodological approach was used. The first stage focussed on methodological quality/internal validity and determined which randomised controlled trials (RCTs) were included. Outcome measures used within these studies were then categorised according to the recommendations of Bombardier (2000) and Deyo et al. (1998). The second stage assessed the exercise quality of the included RCTs, in accordance with the ACSM Exercise Guidelines (2000) (see Table 1).

3.2. Inclusion criteria

Trials were reviewed from 1990 onwards. The systematic review by van Tulder et al. (2000) indicated that the methodological quality of trials had improved since their initial 1991 review. The following inclusion criteria were used.

- Randomised controlled trials of either pragmatic or explanatory design, these being regarded as the most powerful method of determining cause–effect relationships between phenomena (Davidson and Hillier, 2002; Moher et al., 1999).
- Only trials with at least 10 patients in each group were included, in line with the Royal College of General Practitioners (RCGP) selection criteria (Waddell et al., 1999).
- Exercise (either alone or in combination) as the primary intervention under investigation.
- Male and female patients, between the ages of 16 and 74 years.
- Where trials were written in a language other than English, a full translation of the paper was obtained to ensure that all trials were assessed based on all the information available within the publication.
- Patients with spondylosis/spondylolisthesis were included if the degree of slip was two or less. This condition was included as it is considered to be one of the most obvious manifestations of lumbar instability (O'Sullivan et al., 1998).
- Where duplications of trials were sourced, the trial with the highest methodological quality was included (van Tulder et al., 1997).

Elders et al. (2000) and Waddell (1998) have identified sub-acute LBP as pain lasting between 30 days and 12 weeks, and CLBP as pain greater than 12 weeks. In addition, Bouter et al. (1998) indicate that CLBP may occur in multiple episodes over the year. This review specifically focused on chronic non-specific LBP, with or without radiation to the lower limbs. Trials investigating a sub-acute/chronic or chronic/recurrent patient sample were included based on the following criteria.

Table 1
ACSM (2000) exercise guidelines

Muscle strengthening[a]	Frequency of 2–3 days/week, at least one set of 8–12 repetitions at the 8–12 repetition maximum of each exercise using any type of exercise that can be progressed over time.
Flexibility training	Controlled static stretch held for a given duration, intensity to a position of mild discomfort for 10–30 s, 3–4 repetitions of each stretch on 2 or more days of the week.
Cardiovascular endurance[b]	Intensity 55% HR_{max} (220-age); 20 min continuous or intermittent through day using any mode of aerobic exercise involving major muscle groups for at least 6 weeks.

[a] Dynamic stabilisation exercises were classified under the muscle strengthening category. The 8–12 repetition maximum criteria was omitted in this case as this particular type of exercise is usually of low load.

[b] Trials incorporating aerobic exercise that resulted in a significant decrease in resting heart rate were also classified as meeting the necessary cardiovascular training dose (Bowers and Fox, 1988; Janssen, 1987).

- For a sub-acute/chronic sample, at least 75% of the sample must have been CLBP patients;
- For a chronic/recurrent sample, there must have been at least three LBP episodes in the previous year.

3.3. Exclusion criteria

The following were employed as exclusion criteria:

- Trials using the alternate allocation method of randomisation in accordance with the recommendations of Altman and Bland (1999).
- Trials based upon subjects with possible serious spinal pathology, along with spinal surgery patients, if the surgery had been completed less than 1 year.
- Trials having only their abstracts available.
- Trials including patients with fibromyalgia.
- Trials providing insufficient information on the category of low back pain LBP, or if subjects were asymptomatic.
- All trials of low methodological quality (Koes et al., 1995) given that higher quality trials have the most impact on the overall results of a review (de Vet et al., 1997).

Trials of low methodological quality were assessed to see if trial quality affected treatment outcome.

3.4. Search strategy

The most recent systematic review in this area concluded that exercise might help CLBP patients improve return to work rates, and activities of daily living (van Tulder et al., 2002). However, it remains unclear what types of exercises are best; therefore, trials (experimental and pragmatic) involving any type of exercise were sourced. Literature searches were conducted biweekly between October 1st and December 31st 2002 using MEDLINE—Biomed (1990–2002), Amed (1990–2002), CINAHL (1990–2002), PROQUEST medical library (guided search—Jan 1990–Dec 1998 and Jan 1999–Dec 2002), PEDro—simple search (1990–2002), WEB OF SCIENCE (full search/general search 1990–2002), COCHRANE—Dare and Central Register of Controlled Trials and PUBMED—National Library of Medicine (1990–2002). The following search terms were used: LBP, chronic, CLBP, exercise, exercise therapy, physical therapy, physiotherapy, physical activity, sports medicine, strength, flexibility, clinical trials and randomised controlled trials. The following terms were explored: exercise therapy, clinical trials and physical activity. Hand-searches of relevant journals and reference lists of trials, review articles and meta-analyses were also performed.

3.5. Methodological quality

The best way of selecting high-quality physical therapy trials for a systematic review has not yet been determined. A recent study has indicated the need to develop and validate quality scales specific to physical treatments, as certain scales are more suited to a particular trial design (Colle et al., 2002). The van Tulder methodological quality criteria have been recommended by the Cochrane Collaboration Back Review Group for Spinal Disorders (van Tulder et al., 1997). The use of such a standardised set of methodological quality criteria was to score the trials included in this review to facilitate comparison with previous reviews and enhance the consistency of the results (van Tulder et al., 1997). The authors acknowledge that these guidelines have recently been updated, with the introduction of a new rating scale (van Tulder et al., 2003), however, the trials in this review were methodologically scored prior to the publication of this update. Nonetheless, the QUOROM guidelines (Moher et al., 1999) were followed throughout in order to improve the overall quality of this review and, in doing so, have incorporated some of the updated methodological criteria. The 1997 criteria score trials according to 19 item ratings: 11 items assess internal validity, six assess descriptive quality, and two assess statistical validity. Although it is not possible to blind patients to exercise intervention, the authors chose to award a point to trials clearly stating that patients were given minimal information about the differences between interventions. Dropout acceptability was set at $\leq 10\%$ of the total patient sample. The item 'blinding of care provider' was omitted, as it is inapplicable to exercise interventions. Consequently, 10 internal validity criteria were used, giving a maximum van Tulder score of 18 points. These two adjustments were made in accordance with the methods used by Busch et al. (2001), who reviewed exercise training in patients with fibromyalgia.

van Tulder et al. (1997) considered trials to be of high methodological quality if their internal validity was $\geq 5/10$. In order to take a more comprehensive look at methodological quality, the trials in this review are categorised according to the following criteria.

- High quality trials. At least 70% of the methodological criteria met *plus* internal validity of $\geq 6/10$;
- Medium quality trials. At least 50% of methodological criteria met *plus* internal validity of $\geq 5/10$;
- Low quality trials. At least 50% of methodological criteria met *plus* internal validity $< 5/10$.

The methodological quality of the scored RCTs was independently assessed by two reviewers (SDL and JHG). One author (JHG) was blinded to the authors, institution and journal, the other author (SDL) carried out the literature search. Any disagreement between these authors was resolved with a third reviewer (GDB) using the consensus method.

Table 2 details the 51 RCTs scored for methodological quality.

Table 2
van Tulder methodological quality criteria for $n = 51$ scored RCTs

Trial	Internal validity										Descriptive quality						Statistical validity		
	b1	b2	f	g	h	i	j	l	n	p	a	c	d	k	m1	m2	o	q	Quality
Alaranta et al. (1994)	+	?	+	?	−	−	−	?	+	−	+	?	+	?	+	+	+	+	Low
Bendix et al. (1995)	+	+	?	?	−	+	−	−	+	−	+	−	+	+	+	−	+	+	Low
Bendix et al. (2000)	+	+	?	?	+	−	−	−	+	+	?	−	+	?	+	+	+	+	Medium
Bentsen et al. (1997)	+	+	?	+	+	−	−	−	+	−	+	+	+	−	+	+	+	+	Medium
Bronfort et al. (1996)	+	+	+	+	−	+	+	−	+	+	+	+	+	+	+	+	+	+	High
Callaghan (1994)	+	+	?	?	+	−	−	−	+	−	+	?	+	−	+	−	+	+	Low
Chok et al. (1999)	+	?	+	?	−	−	−	−	+	−	+	+	+	+	+	−	+	+	Low
Descarreaux et al. (2002)	+	+	−	+	−	−	−	+	+	+	+	+	−	−	+	−	+	+	Medium
Deyo et al. (1990)	+	+	+	+	+	+	−	−	+	−	+	+	+	+	+	−	+	+	High
Donchin et al. (1990)	+	?	+	?	−	+	−	?	+	+	+	−	+	?	+	+	+	+	Medium
Elnaggar et al. (1991)	+	−	+	?	−	−	−	−	+	−	+	+	+	+	+	−	+	+	Low
Friedrich et al. (1996)	+	−	+	+	−	+	−	−	+	−	+	+	−	−	+	−	+	−	Low
Friedrich et al. (1998)	+	+	+	−	−	−	−	−	+	−	+	+	+	?	+	+	+	+	Low
Frost et al. (1995)	+	+	+	+	−	+	−	−	+	+	+	+	−	−	+	−	−	+	Medium
Frost et al. (1998)	+	+	+	+	−	+	−	−	+	+	+	+	+	+	+	+	+	+	High
Gundewall et al. (1993)	+	+	+	+	−	−	+	−	+	−	−	+	+	−	−	+	+	−	Medium
Hagen et al. (2000)	+	+	−	?	−	+	−	?	+	−	+	+	−	−	+	+	+	+	Low
Hansen et al. (1993)	+	?	?	+	−	+	−	+	+	−	+	+	+	+	+	+	+	+	Medium
Helewa et al. (1999)	+	+	−	−	+	+	−	−	+	+	+	−	+	−	+	+	+	+	Medium
Hemmila et al. (2002)	+	+	?	?	?	+	−	+	+	+	+	−	+	?	+	+	+	+	Medium
Hildebrandt et al. (2000)	+	+	?	?	−	+	−	−	+	+	+	+	+	?	+	+	+	+	Medium
Johannsen et al. (1995)	+	?	+	+	−	−	−	−	+	−	+	+	+	+	+	−	+	+	Low
Kankaanpaa et al. (1999)	+	+	−	−	−	−	−	−	+	−	+	−	+	−	+	+	+	+	Low
Khalil et al. (1992)	+	?	+	+	−	−	−	+	+	−	+	+	+	+	−	−	+	+	Medium
Klaber Moffett et al. (1986)	+	+	?	?	+	+	−	−	+	−	+	+	−	−	+	−	+	+	Medium
Klaber Moffett et al. (1999)	+	+	+	+	−	?	−	+	+	+	+	−	+	−	+	+	+	+	High
Klein and Eek (1990)	+	+	+	?	+	+	−	?	+	−	+	+	+	+	+	−	+	−	Medium
Kuukkanen and Malkia (1996)	−	−	?	?	−	?	−	−	+	−	+	+	+	−	+	+	+	+	Low
Kuukkanen and Malkia (1998)	−	−	+	?	−	?	−	−	+	−	+	+	+	−	+	+	−	+	Low
Kuukkanen and Malkia (2000)	−	−	+	?	?	−	−	−	+	−	+	+	+	−	+	+	+	+	Low
Lindström et al. (1992)	+	+	?	−	−	−	−	+	+	−	+	+	+	−	+	+	+	+	Low
Ljunggren et al., (1997)	+	+	+	+	?	−	−	−	+	−	?	−	+	−	+	+	+	+	Medium
Manniche et al. (1988)	+	+	?	?	−	+	−	−	+	+	+	?	+	+	+	−	−	+	Medium
Manniche et al., (1991)	+	+	?	+	−	+	−	−	+	+	+	?	+	+	+	+	−	+	High
Mannion et al., (1999)	+	+	+	+	+	+	+	−	+	+	+	+	+	−	+	−	+	+	High
Mannion et al., (2001) (a)	+	+	?	?	+	?	−	−	+	−	−	?	?	−	+	−	−	+	Low
Mannion et al., (2001) (b)	+	+	+	?	+	−	−	−	+	+	+	+	+	−	+	+	+	+	Medium
Mellin et al. (1990)	+	?	?	?	−	−	−	−	−	−	?	?	+	−	−	+	+	+	Low
McIlveen and Robertson, (1998)	+	+	+	+	−	+	−	−	+	−	+	+	+	−	+	−	+	+	Medium
O'Sullivan et al. (1997)	+	+	+	+	+	+	−	+	+	−	+	+	+	−	+	+	+	+	High
O'Sullivan et al., (1998)	+	+	+	+	−	+	−	+	+	−	+	+	+	−	+	−	+	+	High
Petersen et al. (2002)	+	−	+	?	−	−	−	−	+	+	+	−	+	?	+	+	+	+	Low
Reilly et al. (1989)	+	?	+	−	−	−	−	+	+	−	−	+	+	−	+	−	+	+	Low
Risch et al., (1993)	+	+	+	+	+	−	−	−	+	−	+	+	+	−	+	−	+	+	Medium
Rittweger et al., (2002)	+	−	+	+	−	−	−	−	+	−	+	−	+	−	+	−	+	+	Low
Sachs et al. (1994)	+	−	−	+	−	−	−	+	+	−	?	?	+	−	+	−	+	+	Low
Snook et al. (1998)	+	+	−	+	+	−	−	−	?	+	+	+	+	+	+	+	+	+	High
Soukup et al. (1999)	+	−	+	+	−	−	−	+	+	−	+	+	+	?	+	+	+	+	Medium
Tortensen et al., (1998)	+	−	+	+	−	+	+	−	+	+	+	+	+	?	+	+	+	+	High
Tritilanunt and Wajanavisit (2001)	+	+	?	+	−	−	−	+	+	−	+	+	+	−	+	−	+	+	Medium
Turner et al. (1990)	+	−	?	?	−	−	−	−	+	−	+	+	+	−	+	+	+	+	Low

(a) Were eligibility criteria specified? (b1) Was the method of randomisation performed? (b2) Was the treatment allocation concealed? (c) Groups similar at baseline (most impt prognostic indicators)? (d) Were index and control interventions explicitly described? (e) Was the care provider blinded to the intervention? (f) Were the co-interventions avoided or comparable? (g) Was compliance acceptable in all groups? (h) Patient blinded to intervention? (i) Outcome assessor blinded to intervention? (j) Relevant outcome measures used? (k) Adverse effects described? (l) Withdrawal/drop out rate described and acceptable? (m1) Was there short-term follow-up measurement? (m2) Was there long-term follow-up measurement? (n) Was the timing of outcome assessment in both groups comparable? (o) Was the sample size in each group described? (p) Did analysis included intention-to-treat analysis? (q) Were point estimates and measures of variability presented for primary OMs? (+) Criterion achieved; (−) criterion not achieved; and (?) insufficient description to decide.

3.6. Strategy for categorisation of exercise

Each trial was initially categorised according to the *main* types of exercise included. This was problematic, given the wide variety of exercise types used within each trial. Consequently, some of the categories overlap. Such a situation makes it difficult to understand the physiological mechanism of action that may be leading to therapeutic improvement (Khalil et al., 1992). It is for the purpose of this review, and to reflect the variety of exercise types offered within programmes that the following broad categories were used.

- *Strengthening/flexibility*: primarily strengthening and muscular endurance exercises for the trunk, upper and/or lower limbs using progressive resistance exercise or the patients body weight as resistance; stretching and range of movement exercises were included to a lesser extent.
- *Aerobic/strengthening*: aerobic exercise lasting longer than 10 min plus muscular endurance and strengthening of trunk, upper, and/or lower limbs as above.
- *Aerobic*: aerobic exercise plus postural instructions.
- *Multimodal*: any functional restoration programme involving behavioural support/positive coping strategies, aerobic and/or strengthening exercise, flexibility/stretching/relaxation, work-hardening, ergonomic advice.
- *Hydrotherapy*: any hydrotherapy-based intervention.
- *Other exercise*: other exercise not stipulated above e.g.: McKenzie; Williams; callisthenics/strengthening; strengthening/ergonomic advice.

Exercise programmes were largely group-based, however, patients were often required to continue their exercises independently between treatment sessions.

3.7. Quality of exercise

The quality of exercise used in trials was assessed using the ACSM guidelines for exercise dose (2000) (Table 1). This was complicated by the variety of exercise used within a single programme (as above), and because the programmes were not always sufficiently detailed. In most cases only the predominant exercise was described. Therefore, the exercise quality of each trial was determined based on the predominant exercise used. Those fulfilling the ACSM (2000) criteria for muscle strengthening, flexibility or cardiovascular endurance, were classified as being of 'high quality exercise'. Those not meeting the criteria were of 'low exercise quality'. Exercise quality was considered 'unclear' if the exercise dose was not specified, or the authors provided insufficient information to reach a score. The quality of McKenzie-based treatment and stabilisation exercises is not accounted for by the ACSM guidelines. Consequently, this presents a limitation to the exercise quality rating. Stabilisation exercises were frequently used in trials as the predominant exercise. For this reason, the authors chose to include those trials using dynamic stabilisation as their predominant exercise within the strengthening category. Table 1 indicates how the guidelines were amended to include such exercise.

3.8. Suitability of outcome measures

No established clinical indicator is generally accepted as a suitable single outcome measure for subjects with LBP (McIlveen and Robertson, 1998). Psychosocial as well as physical aspects of LBP are important in its assessment (Hope, 2002; Staal et al., 2002). In accordance with the World Health Organisation (WHO) International Classification of Functioning and Disability and Health (WHO, 2000), the health of an individual is based on the categories of impairment, activity (previous disability) and participation (previous handicap). Given that all these categories can be influenced by CLBP, the use of outcome measures must adequately reflect the effects and influences that treatment programmes may have on all areas of the patient's life.

Trials were considered as having relevant outcome measures if they included three or more of the five categories of measure recommended by Bombardier (2000) and Deyo et al. (1998). The ICIDH-2 classification is reflected in these five categories as detailed.

- *Back specific function*: Roland–Morris/Oswestry. These instruments reflect the level of activity limitation that patients experience as a result of their back pain.
- *Generic health status*: SF-36/SF-12/EuroQoL. This gives a more comprehensive assessment of the patient's health status than 'back-specific' instruments and can reflect the overall impact of the patient's health status (including co-morbidities) on their role in society.
- *Pain*: frequency and severity of LBP/Body Pain Scale of SF-36/Chronic Pain Grade (optional). This measures impairment but also gives an indication of the extent to which pain interferes with the patient's activities.
- *Work disability*: days off work/days of cut-down work/work status/time to return to work. This reflects the extent to which the patients' condition has a negative influence on their usual role in society.
- *Satisfaction with care/treatment outcome*: Patient Satisfaction Scale/Global question on overall satisfaction. This is considered important in relation to how the patient responds to treatment (Liebenson and Yoemans, 1997). Therefore, it can help improve communication between patient and therapist.

The McGill Pain Questionnaire, Aberdeen Back Pain scale, Sickness Impact Profile and Waddell Disability Index were included in the criteria for acceptability for pain and back-specific function, as there is evidence to support their reliability, validity and responsiveness (Kopec and Esdaile, 1995).

An intervention was considered positive if there was a significant difference ($P < 0.05$), pre- versus post-treatment, in one of the recommended categories of outcome measure. A significant difference in two outcome measures was stipulated in those trials using none of the recommended categories.

3.9. Supervision, compliance and follow-up

An intervention was 'fully supervised' if sessions involved the patient attending a scheduled therapist-led session. 'Partial supervision' involved initial instruction in exercises to complete independently, with periodic follow-up from the therapist to adjust the programme accordingly. The intervention was considered 'unsupervised' if the patient was given a home exercise programme to continue independently, with follow-up only to assess outcome. If no details were given regarding supervision, level of supervision was categorised as 'unclear.'

The trial was considered to have incorporated adequate exercise compliance procedures provided the frequency and type of exercise were recorded in each session, or if patients were asked to keep a diary of exercises completed independently.

Follow-up between 1 and 8 months after the start of treatment was considered short-term. Any follow-up greater than or equal to 9 months after the start of treatment was considered long-term. These end-points were reached to fairly reflect the variations in follow-up between trials, ranging from 2 weeks to 3 years. The cut-off point for short-term follow-up was set at less than 9 months. Since the median treatment period lasted 10 weeks, this cut-off allowed those trials carrying out a 6-month follow-up, starting at the end of treatment, to be included. In addition, the authors considered it unfair to equate a 6-month follow-up with 1 of 2 or 3 years.

4. Results

Fig. 1 identifies the procedure adopted to identify suitable high, and medium, quality trials for this review. A total of 1127 articles was generated by database and hand-searching methods. Eighty-three studies investigated exercise and LBP; five were systematic reviews and 1039 articles did not specifically address the subject under review but facilitated the collection of background information on LBP. Despite the given search criteria, three studies investigated acute low back pain (ALBP) and exercise; these studies were also excluded. Therefore, 80 studies investigated the use of exercise in the management of sub-acute and chronic or recurrent LBP, of which 54 were RCTs. Three of these were excluded prior to methodological scoring as only their abstracts were available. Therefore, 51 RCTs were scored for methodological quality; these trials investigated patients with sub-acute and chronic/recurrent LBP. Of all the RCTs scored the main shortcomings in methodological quality were

- lack of patient blinding,
- lack of relevant outcome measures,
- unacceptable drop-out,
- no intention-to-treat analysis.

Trials of low methodological quality scored poorly on concealment of treatment allocation and adequate compliance (see Table 2).

In accordance with the set criteria, 21 trials were of 'low', 20 were of 'medium' and 10 were of 'high' methodological quality. The 'low' quality trials were further excluded along with 14 'medium'/'high' quality trials: these trials and their reasons for exclusion are available from the authors. Table 3 details the characteristics of included trials. One trial included spondylosis/spondylolisthesis patients with a maximum of 2° of slip (O'Sullivan et al., 1997); this was determined from lateral view radiographs using the method described by Meyerding (1932). These patients were medically suitable to complete an exercise programme. Sixteen RCTs of 'medium' and 'high' methodological quality met all inclusion criteria and were included in this systematic review. All trials had an internal validity score of 50% or more, and all performed an acceptable method of randomisation (see Table 2). Nine of the 16 included trials (56%) reported a positive result in the experimental (exercise) versus the control group. The remaining seven (44%) reported a positive result, but there was no difference between the experimental and control groups: these trials all incorporated exercise into their control treatment. In comparison, 17 (81%) of the 21 'low' quality trials (81%) reported a positive result in the experimental versus the control group. Therefore, inclusion of 'low' quality trials in this review may have had the potential to over-inflate positive results and cause bias (Assendelft et al., 1995; de Bie, 1996; Greenhalgh, 1997; Moher et al., 1999).

4.1. Description of included studies

A total of 22 trials included in the van Tulder et al. (2000, 2002) systematic reviews on exercise and LBP were excluded from this review (Buswell, 1982; Cherkin et al., 1998; Coxhead et al., 1981; Davies et al., 1979; Delitto et al., 1993; Evans et al., 1987; Faas et al., 1995; Farrell and Twomey, 1982; Kendall and Jenkins, 1968; Lidström and Zachrisson, 1970; Lindström et al., 1992; Malmivaara et al., 1995; Manniche et al., 1993; Martin et al., 1980; Nwuga, 1982; Nwuga and Nwaga, 1985; Seferlis et al., 1998; Stankovic and Johnell, 1990; Underwood and Morgan, 1998; Waterworth and Hunter, 1985; White, 1966; Zylbergold and Piper, 1981): these trials investigated ALBP and/or subacute LBP, did not report the duration and type of symptoms, investigated patients post surgery or

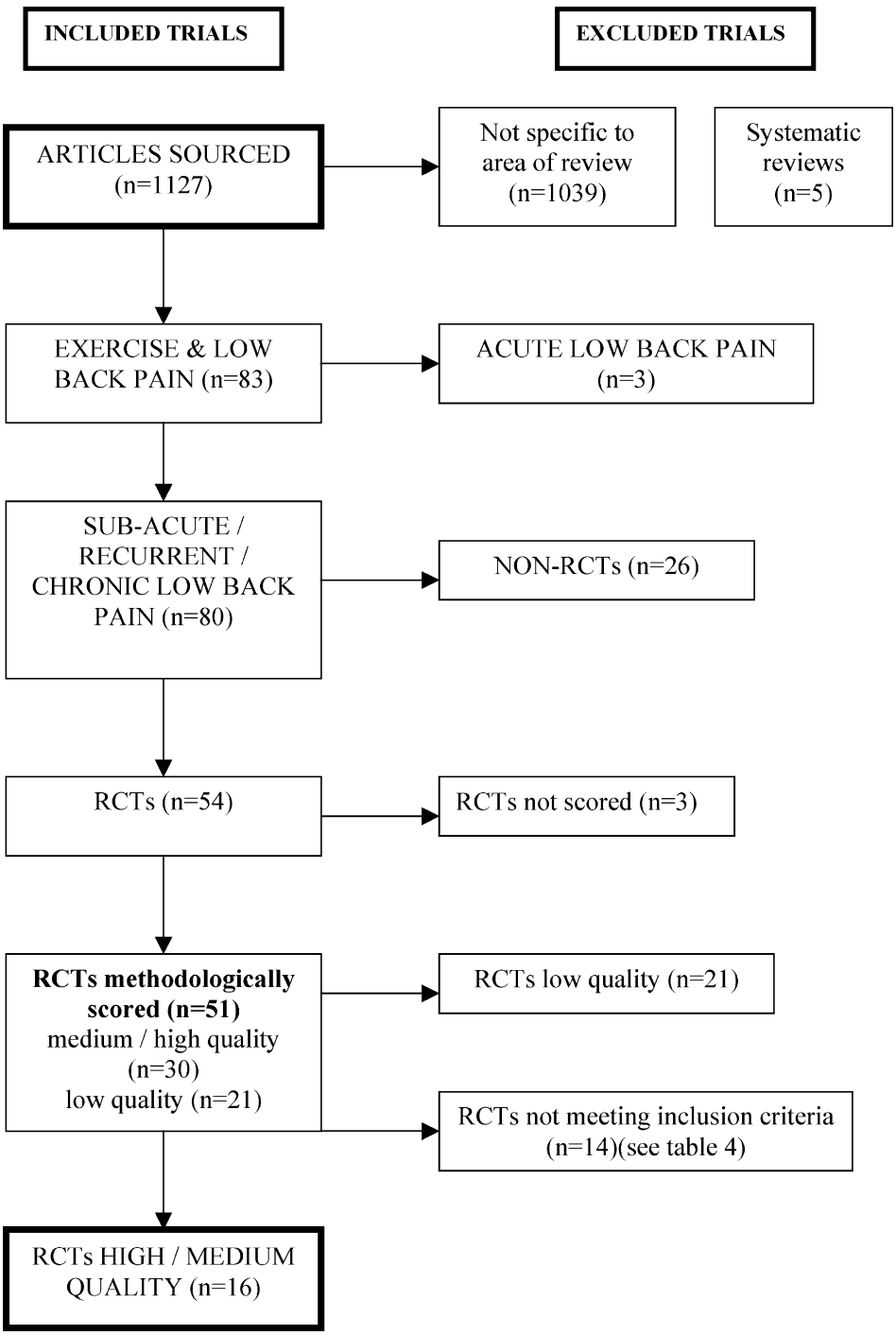

Fig. 1. Flow chart of literature search for chronic low back pain and exercise.

were published before 1990. Seven new trials were included in this review (Bendix et al., 2000; Donchin et al., 1990; Hildebrandt et al., 2000; Khalil et al., 1992; Mannion et al., 1999; McIlveen and Robertson, 1998; Tritilanunt and Wajanavisit, 2001). Table 3 classifies all 16 included trials according to their methodological and exercise quality. Twelve were based upon 'high' quality exercise interventions; of those remaining, three were of 'low' exercise quality and one was 'unclear' due to insufficient information.

4.2. Participants

This review included a total of 1730 males and females between the ages of 16 and 74 years; one trial exclusively used 57-year-old women (Bentsen et al., 1997). The median

Table 3
Included RCTs ($n = 16$) of chronic low back pain and exercise

Study	Type of programme	Methodological quality	Quality of exercise	Acceptable compliance (Y/N)	Outcome measures	Rx outcome	Follow-up (Y/N)	ST/LT follow-up	Outcome maintained (Y/N)
Bendix et al. (2000)	M/MODAL	Medium	High	No	PAIN/WK/DISAB	Aer/Stgth as effective as M/MODAL and cheaper	Yes	ST at LT	Yes (for both gps)
Bentsen et al. (1997)	STGTH/FLEX	Medium	High	Yes	WK DISAB	Supervised gp better compliance and LT results	Yes	ST and LT	Yes
Bronfort et al. (1996)	STGTH/FLEX	High	Unclear	Yes	GEN H/S/RMDQ/ WK DISAB	All gps positive result	Yes	ST and LT	Yes (especially initially supervised programme)
Deyo et al. (1990)	STGTH/FLEX	High	High	Yes	PAIN/B/Spec Func/	Positive effects lost at 2 months	Yes	ST	No
Donchin et al. (1990)	CAL/STGTH	Medium	High	No	OSWESTRY	Positive result maintained at 3 not 6 months	Yes	ST	No
Frost et al. (1998)	M/MODAL	High	High	Yes	OSWESTRY	Positive result maintained at 2 years	Yes	ST and LT	Yes
Hansen et al. (1993)	STGTH/FLEX	Medium	High	Yes	WK DISAB	Positive results lost 1/6/12 months	Yes	ST and LT	No
Hildebrandt et al. (2000)	STGTH/FLEX	Medium	Low	No	None	Positive result maintained ST but not LT	Yes	ST and LT	Yes at 3/6 months not at 12 months
Khalil et al. (1992)	M/MODAL	Medium	Low	Yes	None	Positive result after 2 weeks Rx	No	N/A	N/A
Ljunggren et al. (1997)	STGTH/FLEX	Medium	High	Yes	WK DISAB/Sat. with care	Positive result both gps maintained at 1 year f'up	Yes	ST and LT	Yes
Manniche et al. (1991)	STGTH/FLEX	High	High	Yes	None	Positive result of intensive stgth at 3 months and if c/w ex at 1 year	Yes	ST and LT	Yes (intensive gp c/w ex more and if they did maintained positive results at 1 year f'up)
Mannion et al. (1999)	AER/STGTH	High	High	No	PAIN/RMDQ/Sat. with care	Positive result all gps maintained at 6 months	Yes	ST	Yes
McIlveen and Robertson (1998)	HYDRO	Medium	High	Yes	PAIN/OSWESTRY	Positive result post 4 weeks Rx	Yes	ST	Yes
O'Sullivan et al. (1997)	STGTH/FLEX	High	High	Yes	OSWESTRY	Positive result at 3/6/30 months	Yes	ST and LT	Yes
Risch et al. (1993)	STGTH/FLEX	Medium	Low	Yes	B/Spec Func	Positive result post 10 weeks Rx	Yes	ST	Yes
Tritilanunt and Wajanavisit (2001)	AEROBIC	Medium	High	Yes	None	Positive result of aerobic ex post 12 weeks Rx	Yes	ST	Yes

M/MODAL, multimodal; STGTH, strength; FLEX, flexibility; AER, aerobic; HYDRO, hydrotherapy; CAL, calisthenics; c/w, continue with; WK DISAB, work disability; f'up, follow-up; GEN H/S, general health status; Rx, treatment; RMDQ, Roland Morris Disability Questionnaire; ex's, exercises; B/Spec Func, back specific function; gp (s), group (s); Sat. with care, satisfaction with care; ST, short-term; LT, long-term.

sample size of the control and intervention groups across trials was 42 and 43 patients, respectively; the median size of the smallest group in each trial was 40 patients (range 14–110).

4.3. Interventions

Exercise resulted in positive outcomes in all 16 included RCTs, despite the variety of exercise used within programmes. The most common type of programme was in the category of strengthening/flexibility ($n = 9$). Table 4 details the outcome measures upon which the positive results of exercise were based, and the level of significance reported. Back specific function ($n = 7$ trials) and work disability ($n = 5$ trials) were the most commonly used of the recommended outcome measures to indicate a positive result. Five of the seven trials measuring back specific function reported a positive result in the experimental versus the control group. All five trials assessing work disability showed no difference between the experimental and control groups. Table 4 indicates that the type of control intervention may have had an influence on trial result. Eight out of nine trials reporting a positive effect of exercise in the experimental group used either a waiting list, advice or electrotherapy as the control treatment. The control groups of the trials reporting no difference between groups more commonly used an exercise-based control group. Strengthening was the most popular exercise intervention, as shown in Fig. 2. Fig. 2a shows that four of the strengthening trials targeted the lumbar spine; this figure also highlights that abdominal strengthening was often incorporated with strengthening of the lumbar spine, but was used independently in only one trial. Of the 12 strengthening trials, nine were of 'high' exercise quality. Eight included an element of flexibility, one was part of a multimodal programme, one was part of a hydrotherapy programme, one included aerobic training, and one incorporated strengthening with callisthenics. Therefore, the effects of exercise co-interventions must not be overlooked when interpreting these results. No trials used McKenzie or Williams-based treatment as their predominant exercise.

Fifteen trials were supervised; 11 of these were fully supervised, four were partially supervised. Nine trials carried out both short (<9 months) and long-term (≥ 9 months) follow-up. Of those remaining, six carried out short-term follow-up. Twelve studies maintained their positive results at follow-up, seven of these reported such results at both short-term and long-term follow-up (see Table 3). Table 4 highlights the wide variations in the use of a control/comparison group. Three trials used a waiting list, two used GP advice or no treatment, and the remainder used some other type of exercise as control. This ranged from a home exercise programme to an outpatient programme of a lower intensity than the intervention group.

There was an observed difference in the reporting of exercise compliance between trials of 'high' and 'medium' methodological quality and those of 'low' quality. Seventeen out of 21 (65%) 'low' quality trials either did not monitor or provided insufficient information on compliance. This is in stark contrast to four out of 16 (25%) 'high' and 'medium' quality trials. Bentsen et al. (1997) highlights the positive contribution that supervision of programmes makes to compliance; this is supported in the ACSM (2000) guidelines for exercise prescription. In this review, the trials by Manniche et al. (1991), Bronfort et al. (1996) and Ljunggren et al. (1997) highlighted that patients need to continue with exercise to maintain its positive effects.

A lack of consensus on outcome measurement was evident throughout both included and excluded trials. When not used as a composite score, pain was the most popular outcome measure. Table 3 details the use of the five specific categories of outcome measure within *included* trials. Back specific function was most popular ($n = 8$) followed by work disability ($n = 5$) and pain frequency and intensity ($n = 4$). Generic health status was included in only one trial. Despite the fact that a median of four outcome measures were used, in total, within each included trial, only two trials used three of the recommended categories of outcome measure together. Four trials used none of the recommended categories. Table 4 details the other outcome measures used within trials. Despite the involvement of psychological factors (behavioural, emotional and cognitive) in CLBP (Simmonds et al., 1996; Twomey and Taylor, 2000, p. 352, 355, 356; Waddell et al., 1993) outcome measures reflecting these factors were only included in three of the included trials. These measures included the Community Epidemiologic Scale-Depression (CES-D) (Bronfort et al., 1996), the Mental Health Inventory (Risch et al., 1993), and Objectives Locus of Control (LOC) (Risch et al., 1993). No studies used the Fear Avoidance Beliefs Questionnaire (FABQ) (Waddell et al., 1993), a validated tool focusing specifically on patients' beliefs about how physical activity and work affects their LBP. Twomey and Taylor (2000, p. 364) have indicated that patients who demonstrate fear avoidance beliefs about activity are less likely to comply with a physical exercise programme.

4.4. Adverse effects

Adverse effects were described in six of the 16 RCTs. These ranged from a coronary occlusion (Hansen et al., 1993) and myocardial infarction (Bronfort et al., 1996), both of which were reported as being unrelated to their respective treatment programmes, to an increase in back pain following the start of treatment (Manniche et al., 1991). Of the total trials scored ($n = 51$), 29 did not report adverse effects and nine gave insufficient information to reach a score (see Table 2). It is difficult to establish from these results whether exercise programmes cause adverse effects with CLBP patients.

Table 4
Included RCTs ($n = 16$) result of exercise and level of significance

Study	Number of patients	Predominant exercise	Comparison/control group	Level of supervision	Rx outcome	Recommended outcome measures used	Other outcome measures used	Outcomes on which the positive result was based	Level of significance (P)
Bendix et al. (2000)	127	Aer/stgth/flex	Aer/stgth	Fully supervised	Both groups positive result	PAIN/WK DISAB	ADL assessment; patient's quality of life	WK DISAB	< 0.05
Bentsen et al. (1997)	74	Dynamic stabilisation exercises	HEP	Partial supervision	Both groups positive result	WK DISAB	Schober test; SLR; Million q'aire	WK DISAB	< 0.05
Bronfort et al. (1996)	174	Maximal endurance trunk strengthening	Flex or NSAIDs	Fully supervised	All groups positive result	GEN H/S/RMDQ/ WK DISAB	Schober test; SLR; CES-D; Dartmouth COOP	B/Spec Func WK DISAB	< 0.05
Deyo et al. (1990)	145	FLEX	TENS	Partial supervision	Positive result experimental group	PAIN/B/Spec Func	Activity level; SLR; Schober test	B/Spec Func PAIN	< 0.05
Donchin et al. (1990)	142	STGTH	Waiting list	Fully supervised	Positive result experimental group	OSWESTRY	SLR; Schober test; isometric spinal extensor stgth; abd muscle stgth	B/Spec Func	< 0.001
Frost et al. (1998)	62	STGTH	HEP and advice	Fully supervised	Positive result experimental group	OSWESTRY	None	B/Spec Func	< 0.05
Hansen et al. (1993)	150	Dynamic stabilisation exercises	PT abd ex's or passive PT	Fully supervised	All groups positive result	WK DISAB	VAS—pain intensity; patient's perceived Rx effect	WK DISAB	< 0.05
Hildebrandt et al. (2000)	222	STGTH (Cesar therapy)	Physician's advice	Fully supervised	Positive result experimental group	None	Back pain improvt; P improvt	PAIN Objective posture assessment	< 0.05
Khalil et al. (1992)	28	FLEX (systematic stretching)	FLEX (non-systematic stretching)	Fully supervised	All groups positive result	None	VAS; Back ext stgth; Lsp ROM	PAIN Bk ext stgth	< 0.001
Ljunggren et al. (1997)	103	STGTH (Terapimaster)	STGTH (conventional exercises)	Unsupervised	All groups positive result	WK DISAB/Sat. with care	Ex duration/week	Wk DISAB Sat. with care	< 0.001
Manniche et al. (1991)	90	STGTH (intensive endurance)	Abd and Lsp isometric contractions	Fully supervised	Positive result experimental group	None	Manniche's LBP scale; pain; Schober test	PAIN improvt in daily activities	< 0.05
Mannion et al. (1999)	148	STGTH (progressive resistance)	PT (HEP and advice)	Fully supervised	All groups positive result	PAIN/RMDQ/Sat. with care	Lsp ROM; Coping Strategies Questionnaire	B/Spec Func PAIN	< 0.05
McIlveen and Robertson (1998)	95	STGTH	Waiting list	Fully supervised	Positive result experimental group	PAIN OSWESTRY	SLR; Schober test; lower limb stgth	B/Spec Func	< 0.05

(*continued on next page*)

Table 4 (*continued*)

Study	Number of patients	Predominant exercise	Comparison/control group	Level of supervision	Rx outcome	Recommended outcome measures used	Other outcome measures used	Outcomes on which the positive result was based	Level of significance (*P*)
O'Sullivan et al. (1997)	44	Deep abdominal stabilisation exercises	Physician's advice (regular walks etc.)	Partial supervision	Positive result experimental group	OSWESTRY	VAS—pain intensity; spinal ROM; abdominal recruitment patterns	B/Spec Func PAIN	$P < 0.05$
Risch et al. (1993)	54	Progressive resistance strengthening	Waiting list	Fully supervised	Positive result experimental group	B/Spec Func	Lsp isometric stgth/endurance; activity questionnaire; Mental Health Inventory	PAIN Stgth gains	$P < 0.05$
Tritilanunt and Wajanavisit (2001)	72	AEROBIC	Lsp flexion ex's and advice	Partial supervision	Positive result of aerobic exercise	None	VAS (pain intensity); HR rest; HDL-C	PAIN HR rest decrease	$P < 0.05$

NSAIDs, non-steroidal anti-inflammatory drugs; STGTH, strength; Lsp, lumbar spine; FLEX, flexibility; Abd, abdominal; AER, aerobic; PT, physiotherapy; WK DISAB, work disability; GEN H/S, general health status; Rx, treatment; RMDQ, Roland Morris Disability Questionnaire; ex's, exercises; P Improvt, posture improvement; TENS, transcutaneous electrical neuromuscular stimulation; Bk ext stgth, back extensor strength; HR rest, resting heart rate; Back pain improvt, back pain improvement; ROM, range of motion; B/Spec Func, back specific function; HDL-C, high density lipoprotein cholesterol; LBP, low back pain; Sat. with care, satisfaction with care; HEP, home exercise programme.

Fig. 2. (a) Predominant exercise offered within the included RCTs of chronic low back pain and exercise ($n = 16$). (b) Type of strengthening exercise.

5. Discussion

The results of this systematic review are based on 16 RCTs in which exercise was the primary intervention for CLBP patients. All included trials were of high and medium methodological quality. This indicates that the quality of studies in this review is superior when compared to previous reviews including low quality trials. Twelve included trials were of high exercise quality; eight were of pragmatic and eight of experimental design. Eleven (69%) of included trials compared interventions with differing exercise content and mode of delivery, possibly reflecting the ethical constraints in delaying treatment in any control group.

This review has demonstrated a level of evidence similar to that underpinning existing guidelines for ALBP. The Royal College of General Practitioners (RCGP) Clinical Guidelines for the management of ALBP (Waddell et al., 1999) based their recommendations on the value of advice on eight trials from 1984 to 1995; lack of support for bed rest is based on nine trials from 1961 to 1995. In comparison, all 16 RCTs included in this review demonstrated that, despite the variety of exercise programmes offered, CLBP patients achieve positive results following

exercise therapy. These results were maintained in 12 trials, with supervision a common feature of trials maintaining their positive results at short or long-term follow-up.

5.1. Outcome measures

Dionne et al. (1999) indicate the importance of assessing pain, function and work status in measuring the efficacy of back pain treatments. Measures of impairment often reflect human performance parameters, therefore there is some justification for assessing relevant impairment outcomes in exercise-based trials. However, pain intensity and other measures of impairment, such as the straight leg raise and spinal extensor muscle strength alone are not generally considered suitable as a means of assessing treatment outcome (Beattie and Maher, 1997; Frost et al., 2000). A patient may substantially improve functionally but show little or no change in their level of impairment (Beattie and Maher, 1997; Seeger, 2001; Waddell et al., 1992). Several authors have reported the value of a composite score of pain frequency and intensity (Bombardier, 2000; Deyo et al., 1998; Jensen and McFarland, 1993; Linton et al., 1998). The composite measurement of pain frequency and intensity, as specified by Bombardier (2000) and Linton et al. (1998), is less commonly used within trials. Back specific function ($n = 7$ trials) and work disability ($n = 5$ trials) were the most commonly used of the recommended outcome measures to indicate a positive result. Generic health status, particularly important in populations with co-morbidities (Bombardier, 2000), was assessed in only one trial. A median of four outcome measures was used in each included trial: a median of one outcome measure fell into any of Bombardier's recommended categories. Therefore, the use of a core set of outcome measures, recommended for spinal disorders (Bombardier, 2000), does not appear to have been given priority.

It seems that investigators still focus more on impairment, and not enough on the restoration of activity (previously disability) and participation (previously handicap) despite pain (Manniche, 1996; Mannion et al., 1999). The use of outcome measures must adequately represent the biopsychosocial influences on a patient's recovery. Not only can anxiety increase pain perception and contribute to ineffective coping behaviours (Adams et al., 1994; Simmonds et al., 1996), but significant others may also contribute to the maintenance of illness behaviour (Cohen and Rainville, 2002; Rainville et al., 1997). Patient satisfaction is also very important in relation to how the patient responds to treatment (Liebenson and Yoemans, 1997), yet it is rarely investigated within the trials included in this review. The effect of exercise on the psychology of the individual is beyond the scope of this review. However, it is important to acknowledge the part that psychological aspects play in maintaining the CLBP experience (Adams et al., 1994; Cook and Hassenkamp, 2000), and their influence on compliance with physical exercise programmes. Indeed, patients who demonstrate fear-avoidance beliefs about activity are less likely to comply with a physical exercise programme (Twomey and Taylor, 2000, p. 364). The use of a screening tool, as described by Haldorsen et al. (2002), may help clinicians to identify these patients, who tend to have a poorer prognosis, and require more extensive multidisciplinary treatment programmes (Haldorsen et al., 2002).

5.2. Type of exercise

The value of strengthening exercise for LBP patients has recently been indicated by Vuori (2001) and the Philadelphia Panel (2001). Strengthening was the predominant exercise in 12 out of 16 trials, two-thirds of which were of 'high' exercise quality. The lumbar spine or lumbar spine and lower limbs were the most commonly targeted body sites. This is in keeping with the conclusions of the study by Rainville et al. (1997), who highlighted the importance of strengthening, especially of the lumbar spine extensors, with CLBP patients. Abdominal strengthening was often incorporated with strengthening of the lumbar spine to facilitate trunk stabilisation (Fig. 2a). At least two-thirds of all strengthening programmes incorporated elements of flexibility into their design. A lack of information made it impossible for the authors to assess the quality and, therefore, the specific training effect of flexibility within these programmes. The inclusion of exercise co-interventions introduces a confounding influence when assessing the effectiveness of exercise programmes. This type of design is particularly commonplace in pragmatic trials. At the moment, the possibility that the combination of exercise, or that any exercise type will work, cannot be excluded. Future research must fully explain all exercise interventions included in the trial in order to adequately compare different programmes (Koes et al., 1995; Protas, 1997).

5.3. Supervision

Bentsen et al. (1997) concluded that the supervised exercise group in their study displayed better exercise compliance and long-term results than the home exercise group receiving no feedback. Fifteen out of the 16 included trials were supervised, either fully ($n = 11$) or partially ($n = 4$), of which five maintained their positive results at short-term follow-up and five at both short and long-term follow-up. It would appear from these results that fully or partially supervised programmes might contribute to the maintenance the exercise benefits with CLBP patients (Kelly, 2002). It has been reported that supervision and regular follow-up enable the therapist to adjust a programme according to the patient's progress (Cohen and Rainville, 2002; Descarreaux et al., 2002).

5.4. Compliance

Supervision is thought to play a part in enhancing exercise compliance (ACSM, 2000, p. 162). Within this review, the trials of Bentsen et al. (1997), Bronfort et al. (1996), Ljunggren et al. (1997) and Manniche et al. (1991) (Table 3) highlight the advantages of supervision and exercise compliance. The observed difference in compliance between trials of 'high' or 'medium' (75% acceptable compliance) and those of 'low' methodological quality (15% acceptable compliance) may indicate that 'low' quality trials do not value the importance of compliance levels and consequently do not report them.

The results of this review suggest that supervised trunk strengthening or stabilisation exercises, incorporating flexibility, improve back specific function more than a waiting list, TENS, advice to take regular walks or to independently continue a home exercise programme. However, when both experimental and control groups are given supervised exercise programmes of variable content, both groups achieve a positive result. Perhaps the supervision of exercise programmes enhances CLBP prognosis thereby playing an important part in their success. More research is needed to further investigate this possibility. Despite reports that patient prognosis can direct the content and intensity of treatment for chronic pain conditions (Haldorsen et al., 2002), this was not the key determinant of treatment for trials in this review.

A stronger emphasis must be placed on improving functional activities despite pain (Cohen and Rainville, 2002; Davey and Broadbent, 1998; Deyo and Weinstein, 2001; Frost et al., 2000; Lively, 2002; Manniche, 1996; Rainville et al., 1997), given the fact that psychological dysfunction has a stronger relationship to physical limitation than to pain (Beattie and Maher, 1997; Simmonds et al., 1996). The relevance of outcome measures to the specific condition, in this case CLBP, must also be acknowledged when designing clinical trials. It would be beneficial to adjust the measure of return to work to incorporate the daily activities of all sectors of the adult population. The use of discomfort as opposed to pain may offer patients a re-interpretation of the pain experience (Adams et al., 1994).

Exercise-based treatments aim to utilise the benefits of exercise to promote wellness rather than illness behaviour (Cohen and Rainville, 2002; Rainville et al., 1997; Sollner and Doering, 1997). Patients need to understand why, not just what to do, to facilitate empowerment and commitment to change (Pfingsten et al., 1997; Poulter, 1999). They must play an active role in their treatment to obtain optimum benefit (Adams et al., 1994; Bigos et al., 2001; Goodwin and Goodwin, 2000; Lively, 2002; Staal et al., 2002).

Acknowledgements

This study was supported by a research grant from the Department of Education and Learning (Northern Ireland) into the study of low back pain.

References

Adams N, Ravey J, Bell J. Investigation of personality characteristics in chronic low back pain patients attending physiotherapy out-patient departments. Physiotherapy 1994;80:514–9.

Altman D, Bland J. Treatment allocation in controlled trials: why randomise? Br Med J 1999;318:1209.

American College of Sports Medicine (ACSM), ACSM's guidelines for exercise testing and prescription, 6th ed. Philadelphia, PA: Lippincot Williams and Wilkins; 2000.

Andersson G. Epidemiological features of chronic low back pain. Lancet 1999;354:581–5.

Assendelft WJJ, Koes B, Knipschild PG, Bouter LM. The relationship between methodological quality and conclusions in reviews of spinal manipulation. J Am Med Assoc 1995;274:1942–8.

Beattie P, Maher C. The role of functional status questionnaires for low back pain. Aust J Physiother 1997;43:29–38.

Bentsen H, Lindgarde F, Manthorpe R. The effect of dynamic strength back exercise and/or a home training program in 57 year-old women with chronic low back pain: results of a prospective randomised study with a 3-year follow-up. Spine 1997;22:1494–500.

de Bie RA. Methodology of a systematic review. Phys Ther Rev 1996;1:47–51.

Bigos SJ, McKee J, Holland JP, Holland CL, Hildebrandt J. Back pain, the uncomfortable truth—assurance and activity paradigm. Der Schmertz 2001;15:430–4.

Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders. Spine 2000;25:3100–3.

Bouter LM, van Tulder M, Koes BW. Methodologic issues in low back pain research in primary care. Spine 1998;23:2014–20.

Bowers RW, Fox E. Sports physiology, 3rd ed. Iowa: William C. Brown; 1988.

Bronfort G, Goldsmith C, Nelson CF, Boline PD, Anderson AV. Trunk exercise combined with spinal manipulative or NSAID therapy for chronic low back pain: a randomized, observer-blinded clinical trial. J Manipulative Physiol Ther 1996;19:570–82.

Busch AJ, Schachter CL, Peloso PM. Fibromyalgia and exercise training: a systematic review of randomized clinical trials. Phys Ther Rev 2001;6:287–306.

Buswell J. Low back pain: a comparison of two treatment programmes. N Z J Physiother 1982;10:13–17.

Carpenter DM, Nelson B. Low back strengthening for the prevention and treatment of low back pain. Med Sci Sports Exerc 1999;31:18–24.

Cherkin DC, Deyo RA, Battie M, Street J, Barlow W. A comparison of physical therapy, chiropractic manipulation, and provision of an educational booklet for the treatment of patients with low back pain. N Engl J Med 1998;339:1021–9.

Clinical Standards Advisory Group (CSAG), Report on back pain. London: HMSO; 1994.

Cohen I, Rainville J. Aggressive exercise as treatment for chronic low back pain. Sports Med 2002;32:75–82.

Colle F, Rannou F, Revel M, Fermanian J, Poiraudeau S. Impact of quality scales on levels of evidence inferred from a systematic review of exercise therapy and low back pain. Arch Phys Med Rehabil 2002;83:1745–52.

Cook FM, Hassenkamp A. Active rehabilitation for chronic low back pain—the patient's perspective. Physiotherapy 2000;86:61–8.

Cottingham JT, Maitland J. A three-paradigm treatment model using soft tissue mobilisation and guided movement-awareness techniques for a patient with chronic low back pain: a case study. J Orthop Sports Phys Ther 1997;26:155–67.

Coxhead CE, Inskip H, Meade TW, North WRS, Troup JDG. Multicentre trial of physiotherapy in the management of sciatic symptoms. Lancet 1981;1:1065–8.

Davey R, Broadbent H. Group rehabilitation for chronic back pain: a pilot study. Br J Ther Rehabil 1998;5:636–42.

Davidson I, Hillier V. Comparison of four methods of allocation for clinical trials with small sample sizes. Physiotherapy 2002;88:722–9.

Davies JR, Gibson T, Tester L. The value of exercises in the treatment of low back pain. Rheumatol Rehabil 1979;18:243–7.

Delitto A, Clbulka MT, Erhard RE, Bowling RW, Tenhula JA. Evidence for use of an extension-mobilisation category in acute low back syndrome: a prescriptive validation pilot study. Phys Ther 1993;73:216–28.

Descarreaux M, Normand M, Laurencelle L, Dugas C. Evaluation of a specific home exercise programme for low back pain. J Manipulative Physiol Ther 2002;25:497–503.

Deyo RA, Weinstein J. Low back pain. N Engl J Med 2001;344:363–70.

Deyo RA, Battie M, Beurskens AJHM, Bombardier C, Croft P, Koes B, Malmivaara A, Roland M, von Korff M, Waddell G. Outcome measures for low back pain research: a proposal for standardised use. Spine 1998; 23:2003–13.

Dionne CE, von Korff M, Koepsell TD, Deyo RA, Barlow WE, Checkoway H. A comparison of pain, functional limitations and work status indices as outcome measures in back pain research. Spine 1999;24:2339–45.

Elders LA, van der Beek A, Burdorf A. Return to work after sickness absence due to back disorders: a systematic review on intervention strategies. Int Arch Occup Environ Health 2000;73:339–48.

Evans C, Gilbert JR, Taylor DW, Hildebrandt A. A randomised controlled trial of flexion exercises, education, and bed rest for patients with acute low back pain. Physiother Can 1987;39:96–101.

Faas A. Exercises: which ones are worth trying, for which patients, and when? Spine 1996;21:2874–9.

Faas A, van Eljk JM, Chavannes AW, Gubbels JW. A randomised trial of exercise therapy in acute low back pain. Spine 1995;20:941–7.

Farrell JP, Twomey LT. Acute low back pain: comparison of two conservative treatment approaches. Med J Aust 1982;1:160–4.

Frost H, Lamb S, Shackleton CH. A functional restoration programme for chronic low back pain: a prospective outcome study. Physiotherapy 2000;86:285–93.

Goodwin R, Goodwin N. An audit into a spinal rehabilitation programme. Br J Ther Rehabil 2000;7:275–81.

Greenhalgh T. How to read a paper: papers that summarise other papers (systematic reviews and meta-analyses). Br Med J 1997;315:672–5.

Haldorsen EMH, Grasdal AL, Skouen JS, Risa AF, Kronholm K, Ursin H. Is there a right treatment for a particular patient group? Comparison of ordinary treatment, light multidisciplinary treatment, and intensive multidisciplinary treatment for long-term sick-listed employess with musculoskeletal pain. Pain 2002;95:49–63.

Hartigan C, Rainville J, Sobel JB, Hipona M. Long-term exercise adherence after intensive rehabilitation for chronic low back pain. Med Sci Sports Exerc 2000;32:551–7.

Helewa A, Goldsmith C, Lee P, Smythe HA, Forwell L. Does strengthening the abdominal muscles prevent low back pain? A randomized controlled trial. J Rheumatol 1999;26:1808–15.

Hildebrandt J, Pfingsten M, Saur P, Jansen J. Prediction of success from a multidisciplinary treatment programme for chronic low back pain. Spine 1997;22:990–1001.

Hope P. Assessment and treatment of patients presenting with low back pain and accompanying psychological distress. An evidence-based approach. Physiotherapy 2002;88:745–50.

Indahl A, Velund L, Reikeraas O. Good prognosis for low back pain when left untampered: a randomised clinical trial. Spine 1995;20:473–7.

Jackson D. How is low back pain managed? Retrospective study of the first 200 patients with low back pain referred to a newly established community-based physiotherapy department. Physiotherapy 2001;87: 573–81.

Janssen P. Training lactate pulse rate. Finland: Polar Electro Oy; 1987.

Jensen MP, McFarland CA. Increasing the reliability and validity of pain intensity measurement in chronic pain patients. Pain 1993;55:195–203.

Johannsen F, Remvig L, Kryger P, Beck P, Warming S, Lybeck K, Dreyer V, Larsen LH. Exercises for chronic low back pain: a clinical trial. J Orthop Sports Phys Ther 1995;22:52–9.

Kelly L. Active back management. Sportex Health 2002;12:25–6.

Kendall PH, Jenkins JM. Exercises for backache: a double-blind controlled trial. Physiotherapy 1968;54:154–7.

Kenny D. Exercise-based rehabilitation for injured workers: programme efficacy and identification of factors predicting programme completion and outcome. Int J Rehabil Res 2000;23:7–17.

Klaber Moffett JA, Chase S, Portek I, Ennis JR. A controlled, prospective study to evaluate the effectiveness of a back school in the relief of chronic low back pain. Spine 1986;11:120–2.

Koes BW, Bouter LM, van der Heijden GJMG. Methodological quality of randomised clinical trials on treatment efficacy in low back pain. Spine 1995;20:228–35.

Kopec JA, Esdaile J. Functional disability scales for back pain. Spine 1995; 20:1943–9.

Lahad A, Malter AD, Berg AO, Deyo RA. The effectiveness of four interventions for the prevention of low back pain. J Am Med Assoc 1994;272:1286–91.

Lidström A, Zachrisson M. Physical therapy on low back pain and sciatica. Scand J Rehabil Med 1970;2:37–42.

Liebenson C, Yoemans S. Outcomes assessment in musculoskeletal medicine. Man Ther 1997;2:67–74.

Lindström I, Őhlund C, Eek C, Wallin L, Peterson LE, Nachemson A. Mobility, strength, and fitness after a graded activity programme for patients with subacute low back pain. A randomised prospective clinical study with a behavioural therapy approach. Spine 1992;17:641–52.

Linton SJ, Hellsing A, Hallden K. A population-based study of spinal pain among 35–45 year old individuals. Prevalence, sick leave and health care use. Spine 1998;23:1457–63.

Lively M. Sports medicine approach to low back pain. South Med J 2002; 95:642–6.

Ljunggren AE, Weber H, Kogstad O, Thom E, Kirkesola G. Effect of exercise on sick leave due to low back pain: a randomized, comparative, long-term study. Spine 1997;22:1610–6.

Malmivaara A, Häkkinen U, Aro T, Heinrichs ML, Koskenniemi L, Kuosma E, Lappi S, Paloheimo R, Servo C, Vaaranen V, Hernberg S. The treatment of acute low back pain—bed rest, exercises or ordinary activity? N Engl J Med 1995;332:351–5.

Manniche C. Clinical benefit of intensive dynamic exercises for low back pain. Scand J Med Sci Sports 1996;6:82–7.

Manniche C, Lundberg E, Christensen I, Bentzen L, Hesselhoe G. Intensive dynamic back exercises for chronic low back pain: a clinical trial. Pain 1991;47:53–63.

Manniche C, Asmussen K, Lauritsen B, Vinterberg H, Karbo H, Ablidstrup S, Fischer-Nielsen K, Krebs R, Ibsen K. Intensive dynamic back exercises with or without hyperextension in chronic back pain after surgery for lumbar disc protrusion: a clinical trial. Spine 1993;18:560–7.

Mannion AF, Muntener M, Taimela S, Dvorak J. A randomised clinical trial of three active therapies for chronic low back pain. Spine 1999;24: 2435–48.

Martin PR, Rose MJ, Nichols PJR, Russell PL, Hughes JG. Physiotherapy exercises for low back pain: process and clinical outcome. Int J Rehabil Med 1980;8:34–8.

McIlveen B, Robertson V. A randomised controlled study of the outcome of hydrotherapy for subjects with low back or back and leg pain. Physiotherapy 1998;84:17–26.

Meyerding H. Spondylolisthesis. Surg, Gynecol Obstet 1932;54:371–7.

Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Lancet 1999;354:1896–900.

Nwuga VCB. Relative therapeutic efficacy of vertebral manipulation and conventional treatment in back pain management. Am J Phys Med 1982;61:273–8.

Nwuga G, Nwaga V. Relative therapeutic efficacy of the Williams and McKenzie protocols in back pain management. Physiother Pract 1985; 1:99–105.

O'Sullivan PB, Twomey L, Allison GT. Altered abdominal muscle recruitment in patients with chronic low back pain following a specific exercise intervention. J Orthop Sports Phys Ther 1998;27:114–24.

Pfingsten M. Functional restoration—it depends on an adequate mixture of treatment. Der Schmertz 2001;15:492–8.

Pfingsten M, Hildebrandt J, Leibing E, Franz C, Saur P. Effectiveness of a multimodal treatment program for chronic low-back pain. Pain 1997; 73:77–85.

Philadelphia Panel, Philadelphia Panel evidence-based clinical practice guidelines on selected rehabilitation interventions for low back pain. Phys Ther 2001;81:1641–74.

Poulter D. Letters to the editor-in-chief; empower the patient. J Orthop Sports Phys Ther 1999;29:616–7.

Protas E. Aerobic exercise in the rehabilitation of individuals with chronic low back pain: a review. Crit Rev Phys Rehabil Med 1997;8:283–95.

Quittan M. Management of back pain: review. Disabil Rehabil 2002;24: 423–34.

Rainville J, Sobel J, Hartigan C, Monlux G, Bean J. Decreasing disability in chronic back pain through aggressive spine rehabilitation. J Rehabil Res Dev 1997;34:383.

Risch SV, Norvell N, Pollock ML, Risch ED, Langer H, Fulton M, Graves JE, Leggett SH. Lumbar strengthening in chronic low back pain patients: physiologic and psychological benefits. Spine 1993;18: 232–8.

Rittweger J, Just K, Kautzsch K, Reeg P, Felsenberg D. Treatment of chronic lower back pain with lumbar extension and whole-body vibration exercise: a randomised controlled trial. Spine 2002;27: 1829–34.

Seeger D. Physiotherapy in low back pain—indications and limits. Der Schmertz 2001;15:461–7.

Seferlis T, Nemeth G, Carlsson AM, Gillstrőm P. Conservative treatment in patients sick-listed for acute low back pain: a prospective randomised study with 12 months follow-up. Eur Spine J 1998;7:461–70.

Simmonds MJ, Kumar S, Lechelt E. Psychosocial factors in disabling low back pain: causes or consequences? Disabil Rehabil 1996;18:161–8.

Sollner W, Doering S. A psychological therapy methods in chronic non-radicular backache: review. Orthopade 1997;26:535–43.

Staal JB, Hlobil H, van Tulder MW, Köke AJA, Smid T, van Mechelen W. Return-to-work interventions for low back pain: a descriptive review of contents and concepts of working mechanisms. Sports Med 2002;32: 251–67.

Stankovic R, Johnell O. Conservative treatment of acute low back pain: a prospective randomised trial. Spine 1990;15:120–3.

Taimela S, Harkapaa K. Strength, mobility, their changes and pain reduction in active functional restoration for chronic low back disorders. J Spinal Disord 1996;9:306–12.

Taimela S, Diederich C, Hubsch M, Heinricy M. The role of physical exercise and inactivity in pain recurrence and absenteeism from work after active outpatient rehabilitation for recurrent or chronic low back pain: a follow-up study. Spine 2000;25:1809–16.

Tortensen TA, Ljunggren A, 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;23:2616–24.

van Tulder M. Treatment of low back pain: myths and facts. Der Schmertz 2001;15:499–503.

van Tulder MW, Koes B, Bouter LM. Conservative treatment of acute and chronic non-specific low back pain: a systematic review of randomised controlled trials of the most common interventions. Spine 1997;22: 2128–56.

van Tulder M, Malmivaara A, Esmail R, Koes B. Exercise therapy for low back pain. A systematic review within the framework of the Cochrane Collaboration Back Review Group. Spine 2000;25: 2784–96.

van Tulder M, Malmivaara A, Esmail R, Koes B. Exercise therapy for low back pain (Cochrane Review). Cochrane Lib 2002;3.

van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. Spine 2003;28:1290–9.

Twomey LT, Taylor J. Physical therapy of the low back, 3rd ed. Philadelphia, PA: Churchill Livingstone; 2000.

Underwood MR, Morgan J. The use of a back class teaching extension exercises in the treatment of acute low back pain in primary care. Fam Pract 1998;15:9–15.

de Vet HC, de Bie R, van der Heijden GJ, Verhagen AP, Sijpkes P, Knipschild PG. Systematic reviews on the basis of methodological criteria. Physiotherapy 1997;83:284–9.

Vuori IM. Dose-response of physical activity and low back pain, osteoarthritis, and osteoporosis. Med Sci Sports Exerc 2001;33(Supplement):S551–86.

Waddell G. The back pain revolution. London: Churchill Livingstone; 1998.

Waddell G, Somerville D, Henderson I, Newton M. Objective clinical evaluation of physical impairment in chronic low back pain. Spine 1992;17:617–28.

Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain 1993; 52:157–68.

Waddell G, McIntosh A, Hutchinson A, Feder G, Lewis M. Low back pain evidence review. London: Royal College of General Practitioners (RCGP); 1999.

Waterworth RF, Hunter IA. An open study of diflunisal, conservative, and manipulative therapy in the management of acute mechanical low back pain. N Z Med J 1985;95:372–5.

White AWM. Low back pain in men receiving workmen's compensation. Can Med Assoc J 1966;95:50–6.

World Health Organisation, 2000. International Classification of Functioning and Disability (ICIDH-2), Geneva: Beta-2 draft, Full version, Strand, 10 April 2000.

Zylbergold RS, Piper MC. Lumbar disc disease: comparative analysis of physical therapy treatments. Arch Phys Med Rehabil 1981;62: 176–9.