SPINE Volume 30, Number 19, pp 2153–2163

Back Schools for Nonspecific Low Back Pain

A Systematic Review Within the Framework of the Cochrane Collaboration Back Review Group

M. W. Heymans, PT, PhD,* M. W. van Tulder, PhD,†‡ R. Esmail, BSc, MSc,§ C. Bombardier, MD, PhD,‖ and B. W. Koes, PhD¶

Study Design. A systematic review within the Cochrane Collaboration Back Review Group.

Objectives. To assess the effectiveness of back schools for patients with nonspecific low back pain (LBP).

Summary of Background Data. Since the introduction of the Swedish back school in 1969, back schools have frequently been used for treating patients with LBP. However, the content of back schools has changed and appears to vary widely today.

Methods. We searched the MEDLINE and EMBASE databases and the Cochrane Central Register of Controlled Trials to November 2004 for relevant trials reported in English, Dutch, French, or German. We also screened references from relevant reviews and included trials. Randomized controlled trials that reported on any type of back school for nonspecific LBP were included. Four reviewers, blinded to authors, institution, and journal, independently extracted the data and assessed the quality of the trials. We set the high-quality level, *a priori*, at a trial meeting six or more of 11 internal validity criteria. Because data were clinically and statistically too heterogeneous to perform a meta-analysis, we used a qualitative review (best evidence synthesis) to summarize the results. The evidence was classified into four levels (strong, moderate, limited, or no evidence), taking into account the methodologic quality of the studies. We also evaluated the clinical relevance of the studies.

Results. Nineteen randomized controlled trials (3,584 patients) were included in this updated review. Overall, the methodologic quality was low, with only six trials considered to be high-quality. It was not possible to perform relevant subgroup analyses for LBP with radiation *versus* LBP without radiation. The results indicate that there is moderate evidence suggesting that back schools have better short- and intermediate-term effects on pain and functional status than other treatments for patients with recurrent and chronic LBP. There is moderate evidence suggesting that back schools for chronic LBP in an occupational setting are more effective than other treatments and placebo or waiting list controls on pain, functional status, and return to work during short- and intermediate-term follow-up. In general, the clinical relevance of the studies was rated as insufficient.

Conclusion. There is moderate evidence suggesting that back schools, in an occupational setting, reduce pain and improve function and return-to-work status, in the short- and intermediate-term, compared with exercises, manipulation, myofascial therapy, advice, placebo, or waiting list controls, for patients with chronic and recurrent LBP. However, future trials should improve methodologic quality and clinical relevance and evaluate the cost-effectiveness of back schools.

Key words: systematic review, back scoliosis, Cochrane Collaboration, low back pain, effectiveness, clinical relevance. **Spine 2005;30:2153–2163**

From the *Department of Public and Occupational Health/EMGO-Institute, VU University Medical Center, Body@Work TNO VUmc, Amsterdam, The Netherlands; †Institute for Research in Extramural Medicine, VU University Medical Center, Amsterdam, The Netherlands; ‡Institute for Health Sciences, Faculty of Earth & Life Sciences, Vrije University, Amsterdam, The Netherlands; §Calgary Regional Health Authority, Calgary, Alberta, Canada; ‖Institute for Work & Health, Toronto, Ontario, Canada; and ¶Department of General Practice, Erasmus Medical Center, Rotterdam, The Netherlands.

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

Foundation funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.

This study is granted by Dutch Health Insurance Board. The funding organizations had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

One of the authors (Claire Bombardier) is coordinating editor of the Cochrane Back Review Group. Editors are required to conduct at least one Cochrane review. This requirement ensures that editors are aware of the processes and commitment needed to conduct reviews. None of the editors are first authors. This involvement does not seem to be a source of conflict of interest in the Back Review Group. Any editor who is a reviewer is excluded from editorial decisions on the review in which they are contributors.

Address correspondence and reprint requests to Martijn Heymans, PT, PhD, Department of Public and Occupational Health/EMGO-Institute, VU University Medical Center, Body@Work TNO Vumc, Van der Boechorststraat 7, Amsterdam, The Netherlands, 1081 BT; E-mail: mw.heymans@vumc.nl

Low back pain (LBP)-related disability and work absence account for high economic costs in Western societies.[1] Direct and indirect costs in the United States were estimated to be more than U.S. $50 billion per year.[2] Estimates of direct and indirect costs in the United Kingdom in 1998 were U.S. $11 billion.[3] Estimates of the financial burden of LBP in the Netherlands in 1991 indicated that the total costs were almost U.S. $5 billion.[4] Although LBP rarely indicates a serious underlying disorder, patients with LBP that lasts for longer than 1 to 2 months have an elevated risk of developing longer-term disability and repeated care-seeking.[5] Moreover, the recovery process of patients with chronic LBP is slow, and their demands on the healthcare system are both large and costly. To date, many therapeutic interventions have been performed and studied for the treatment of LBP; however, no single treatment has proven to be obviously superior compared with any other.[6,7] Consequently, there are discrepancies between countries in the various clinical guidelines and therapeutic recommendations for patients with LBP.[8–10] Continuously and systematically summarizing the literature provides the best evidence for the treatment of (subgroups of) patients with LBP. In this

systematic review, we will present the results on the effectiveness of back schools for nonspecific LBP.

The original "Swedish back school" was introduced by Zachrisson-Forsell in 1969. It was intended to reduce the pain and prevent recurrences of episodes of LBP.[11,12] The back school consisted of information on the anatomy of the back, biomechanics, optimal posture, ergonomics, and back exercises. Four small group sessions were scheduled during a 2-week period, each session lasting 45 minutes. Since the introduction of the Swedish back school, the content and length of back schools have changed and appear to vary widely today.

This review is an update of a previously conducted Cochrane review of randomized controlled trials (RCTs) on the effectiveness of back schools[13] and two systematic reviews on back schools and group education interventions for LBP.[14,15] In these reviews, it was not possible to statistically pool the studies because of the large data deficiencies and heterogeneity in trial designs. Conclusions were generated on the basis of the methodologic quality scores of the studies, assessed using a generally accepted criteria list, in combination with a best evidence synthesis. It was indicated that modifications of the Swedish back school, offering quite an intensive program in an occupational setting, seemed to be the most effective type of back school. However, evidence for the cost-effectiveness of back schools was lacking. Since 1998, five new RCTs have been conducted that evaluated the effectiveness of back schools. In addition, method guidelines for systematic reviews in the field of back pain were recently published that contained new recommendations.[16]

Objectives

The objective of this systematic review was to determine if back schools were more effective than other treatments or no treatment for patients with nonspecific LBP. Criteria for considering studies for this review include the following:

Types of Studies

Only RCTs were included. Nonrandomized trials were excluded.

Types of Participants

Randomized controlled trials that included subjects with nonspecific LBP, 18 to 70 years of age, were included. LBP was defined as pain localized below the scapulas and above the cleft of the buttocks; nonspecific indicated that no specific cause was detectable, such as infection, neoplasm, metastasis, osteoporosis, rheumatoid arthritis, fracture, or inflammatory process.

Types of Interventions

RCTs in which one of the treatments consisted of a back school type of intervention were included. A back school was defined as consisting of an educational and skills acquisition program, including exercises, in which all lessons were given to groups of patients and supervised by a paramedical therapist or medical specialist. Additional interventions were allowed. However, if the back school was part of a larger multidisciplinary treatment program, the study was only included as long as a contrast existed for the back school. For example, a study comparing a fitness program with a back school plus fitness program was included, but a study comparing a back school plus fitness program with a waiting list control group (WLC) not.

Types of Outcome Measures

Randomized controlled trials that measured at least one of the four primary outcome measures that are considered to be the most important for back pain, that is, return to work (return to work status, days off work), pain (VAS), a global measure of improvement (overall improvement, proportion of patients recovered, subjective improvement of symptoms), and functional status (expressed on a back-specific index, such as the Roland Disability Questionnaire or the Oswestry Scale) were included.[17,18] Physiologic outcomes of physical examination, such as range of motion, spinal flexibility, degrees of straight leg raising, or muscle strength were considered secondary outcomes because these outcome measures may correlate poorly with the clinical status of the patient.[19] Other symptoms such as medication use and side effects were also considered.

Search Strategy for Identification of Studies

RCTs published in English, Dutch, French, and German were included. The highly sensitive literature searches in MEDLINE and EMBASE were based on the search strategies recommended and updated by the Editorial Board of the Cochrane Back Review Group[16] and Robinson *et al*.[20] The following search strategy was conducted for the original review. Randomized controlled trials were identified by:

(**A**) a computer-aided search of the MEDLINE (1966–1997) and EMBASE (1988–1997) databases;
(**B**) screening references given in relevant reviews and included RCTs;
(**C**) screening CENTRAL, the Cochrane library 1998, Issue 4, using the search terms "back pain" and "low back pain."

For the updated review, the same searches were conducted in the same databases for the period January 1998 to November 2004, and the Cochrane Library 2004, Issue 4 was used to identify new studies.

Methods of the Review

For the 2004 update, two reviewers (M.H., M.vT.) independently selected new studies, assessed the methodologic quality and extracted the data (using a standardized form). This was conducted in the same way, described in the following sections, as in the previously published systematic review.

Study Selection

Two reviewers independently selected the trials to be included in the systematic review according to the complete search strategy; *i.e.*, they ran the search strategy and selected the RCTs. A consensus method was used to solve disagreements about the selection of RCTs, and a third reviewer was consulted if disagreement persisted. The study selection was completed in two steps. In Step 1, the two reviewers first screened the titles, abstracts, and key words of all references identified by the literature search to determine if they met the inclusion criteria. In Step 2, the full text was retrieved for studies for which the inclusion decision could not be made by screening in Step 1, and reviewed against the inclusion criteria.

Methodologic Quality Assessment

The methodologic quality of the RCTs was independently assessed by two reviewers. A consensus method was used to resolve disagreements and a third reviewer was consulted if disagreements persisted. If the article did not contain information on (one or more of) the methodologic criteria (score "unclear"), the authors were contacted for additional information. We anticipated that authors might work at other places than listed in the publications. We therefore tried to locate their current working address through their last publication in MEDLINE or through the Internet. If we could not find a more recent working address, we sent the request for information to the address listed on the paper used in our review. If the authors could not be contacted or if the information was no longer available, the criteria were scored as "unclear."

Clinical Relevance

Two reviewers independently scored the clinical relevance of the included studies according to five questions recommended by the Cochrane Collaborative Back Review Group.[16] Each question was scored positive (+) if the clinical relevance item was fulfilled, negative (−) if the item was not fulfilled, and unclear (?) if data were not available. The five questions are:

1. Are the patients described in detail so that you can decide whether they are comparable to those that you see in your practice?
2. Are the interventions and treatment settings described well enough so that you can provide the same for your patients?
3. Were all clinically relevant outcomes measured and reported?
4. Is the size of the effect clinically important?
5. Are the likely treatment benefits worth the potential harms?

Data Extraction

Two reviewers independently extracted the data (using a standardized form) on the four types of outcomes: a) pain intensity: expressed on a visual analog (VAS) or similar scale, b) overall improvement: proportion of patients recovered or improved, c) functional status: expressed on a back-pain specific scale (*e.g.*, Roland Disability Questionnaire, Oswestry Questionnaire) or a generic scale (*e.g.*, Sickness Impact Profile), and d) return to work: number of days of sick leave or proportion of patients returned to work.

Analysis

The results of each RCT were plotted as point estimates with corresponding 95% confidence intervals (CI). Statistical homogeneity was formally tested, and the clinical homogeneity was evaluated by exploring the differences between the RCTs, taking into consideration the study population, types of back schools and reference treatments, timing of follow-up measurements, and outcomes and measurement instruments. On the basis of these evaluations, attempts were made to statistically pool the data for the outcome measures pain, functional status, and return to work, for the comparisons back school *versus* other treatments and back school *versus* WLC or placebo. These attempts were made for short-, intermediate-, and long-term follow-up. As shown in Results, many studies did not report their results in a way that enabled us to perform statistical pooling (for example, for continuous data, means were presented but no standard deviations). Consequently, for most comparisons, only a limited number of studies were available for statistical pooling. Furthermore, studies were heterogeneous with respect to study populations, interventions, and settings. Therefore, we did not perform a meta-analysis but summarized the results using a rating system with four levels of evidence (best evidence synthesis), based on the quality and the outcome of the studies:

> Strong evidence—provided by generally consistent findings in multiple high-quality RCTs;
>
> Moderate evidence—provided by generally consistent findings in one high-quality RCT plus one or more low quality RCTs, or by generally consistent findings in multiple low quality RCTs;
>
> Limited or conflicting evidence— only one RCT (either high or low quality) or inconsistent findings in multiple RCTs;
>
> No evidence—no RCTs.

We defined high-quality studies as RCTs that fulfilled six or more of the internal validity criteria. We also performed sensitivity analyses, exploring the results, when high quality was defined as fulfilling five or more or seven or more, or if high quality was defined as having adequate concealment of treatment allocation.

Analyses were conducted separately for: a) (sub)acute LBP (lasting 12 weeks or less) and chronic LBP (lasting longer than 12 weeks) and b) back schools in an occupational setting. It was not possible to make relevant subgroup analyses for LBP with radiation *versus* LBP without radiation. RCTs that included a mixed population of patients with LBP were scored for clinical relevance but were excluded from the analysis. A study was defined as

Table 2. Definitions of Internal Validity Criteria

1a) *Method of randomization* - a random (unpredictable) assignment sequence. Examples of adequate methods are computer generated random number table, and use of sealed opaque envelopes. Methods of allocation using date of birth, date of admission, hospital numbers, or alternation should not be regarded as appropriate.
1b) *Concealment of treatment allocation* - assignment generated by an independent person not responsible for determining the eligibility of the patients. This person has no information about the persons included in the trial and has no influence on the assignment sequence or on the decision about eligibility of the patient.
2) *Withdrawal/dropout rate* - the number of participants who were included in the study but did not complete the observation period or were not included in the analysis must be described and reasons given. If the percentage of withdrawals and dropouts does not exceed 20% for short-term follow-up and 30% for long-term follow-up and does not lead to substantial bias, a yes is scored. (These percentages are arbitrary, not supported by literature.)
3) *Co-interventions avoided or equal* - Co-interventions should either be avoided in the trial design or comparable between the index and control groups.
4) *Blinding of patients* - The reviewer determines if there was enough information about the blinding of the patient to score a yes. Because it is difficult to blind the patients for a back school program, we considered the blinding also adequate if an attempt was made to blind the patients or if the credibility of the treatments was evaluated and treatments were equally credible and acceptable to patients.
5) *Blinding of observer* - The reviewer determines if there was enough information about the blinding of the outcome assessor to score a yes.
6) *Intention-to-treat analysis* - All randomized patients are reported/analyzed in the group to which they were allocated by randomization for the most important moments of effect measurement (minus missing values) irrespective of noncompliance and co-interventions.
7) *Similarity of baseline characteristics* - In order to receive a yes, groups have to be similar at baseline regarding demographic factors, duration and severity of complaints, percentage of patients with neurologic symptoms, and value of main outcome measure(s).
8) *Adequate length of follow-up* - scored positive if an effect measurement is included after 12 months or more.
9) *Blinding of care provider* - The reviewer determines if there was enough information about the blinding of the care provider to score a yes. Becasue it is probably impossible to blind care providers to whether or not they were giving a back school intervention, this item did not apply here.
10) *Compliance* - The reviewer determines if the compliance to the interventions is acceptable, based on the reported intensity, duration, number, and frequency of sessions for both the index intervention(s) and control intervention(s).

being conducted in an occupational setting when the study population consisted of a working population or workers on sick leave.

Description of Studies

We ended up including 19 studies that examined 3,584 patients and were reported in 27 papers (Table 1, available for viewing online through Article Plus only). Four studies included a homogeneous population of LBP patients without radiation,[21–24] while seven studies did not specify if patients had radiating symptoms or not and eight studies included a mixed population of patients with and without radiating symptoms. Of the 19 studies, five studies reported on acute/subacute LBP,[23,25–28] Nineteen studies reported on chronic LBP patients,[21,29–46] and three studies reported on a mixed population of acute and chronic LBP patients.[22,24,47] Three studies did not report any data on the sex and age of the groups evaluated,[30,41,47] in four studies the study population consisted of women only,[35–37,39] and two studies included only men.[42,43] Table 1 (available for viewing online through Article Plus only) shows that the back school interventions varied from a very intensive 3-week inpatient program[31–34] to a "Swedish back school" consisting of four lessons totaling 3 hours.[21] The reference treatments also varied widely from WLC[38–41] to exercise therapy,[48,49] spinal manipulation[29,47] or oral or written instructions.[22,31–37]

Methodologic Quality of Included Studies

The same version of the criteria list of our previously published systematic reviews[13] was used to assess the methodologic quality of the RCTs (Table 2). Compared with the original criteria list, equal weights were assigned to all criteria. The items were scored as positive (+), negative (−), or unclear (?). See Table 2 for operationalization of the criteria.

■ Results

Study Selection

Our original 1998 literature search resulted in the identification of 47 references from MEDLINE and 252 from EMBASE. However, 28 references were included in both MEDLINE and EMBASE, leaving a total of 271. The first selection, based on titles, key words, and abstracts, resulted in 27 disagreements between the two reviewers. After discussing these disagreements, the reviewers decided to include two, exclude 15 abstracts, and they were

Table 3. Characteristics of Excluded Studies

Study	Characteristics
Stankovic[48,49] (1990)	Did not fulfill inclusion criteria. Back school intervention consisted of education only, without exercises.
Molde Hagen[50,51] (2003)	Did not fulfill inclusion criteria. The back school was not applied to groups of patients.
Frost[52,53] (1995)	Fatal error. An appropriate contrast for the back school was not used in their design.
Morrison[54] (1988)	Fatal error. Each group was assessed once: the control group at the beginning of the program and the back school group at the end.
Mucha[55] (1996)	Fatal error. Study population randomized into back school and control group, but difference between groups was not analyzed. Only data on back school group.
Roques[56] (2002)	Did not fulfill inclusion criteria. Patients were not randomly allocated to treatment.

Table 4. Methodologic Criteria

Study	1a	1b	2	3	4	5	6	7	8	9	10	Total Score
Berquist[25]	+	?	+	+	−	?	+	?	+	−	−	5
Berwick[22]	?	?	+	+	−	−	−	+	+	−	−	4
Dalichau[42,43]	?	?	+	?	−	−	−	+	−	−	?	2
Donchin[40]	−	+	+	−	−	−	+	+	+	−	−	5
Herzog[47]	?	?	−	?	−	+	−	?	−	−	−	1
Hsieh[24]	+	+	+	?	−	−	+	+	−	−	+	6
Hurri[35–37]	?	−	+	−	−	−	+	+	+	−	−	4
Härkäpää[31–34]	?	?	+	+	−	−	+	+	+	−	−	5
Indahl[27,28]	−	−	+	−	+	+	+	−	+	−	+	6
Keijsers[38]	?	?	+	−	−	−	?	−	−	−	−	1
Keijsers[41]	?	?	−	?	−	−	−	?	−	−	−	0
Klaber Moffett[29]	+	+	+	−	−	+	+	+	+	−	−	7
Lankhorst[21]	−	−	+	−	−	−	−	+	+	−	?	3
Leclaire[23]	+	?	+	?	−	−	+	+	+	−	+	6
Lindequist[26]	−	−	+	−	−	−	+	−	−	−	−	2
Linton[39]	+	+	+	−	−	+	+	+	−	−	?	6
Lønn[44,45]	+	+	+	−	−	−	−	+	+	−	+	6
Penttinen[46]	?	?	−	−	−	−	−	+	+	−	−	2
Postacchini[30]	?	?	+	?	−	−	−	?	−	−	−	1

not sure about inclusion of 10 abstracts. After consensus, the final results of the first step in study selection were that both reviewers agreed that 14 abstracts met the inclusion criteria, 23 were rated as "not sure" and 234 were excluded. Copies of the full papers of the 23 abstracts rated "not sure" were subsequently assessed, resulting in the inclusion of an additional seven papers. Consequently, a total of 21 papers met our selection criteria. However, the four papers that reported on one trial,[31–34] and the three papers that reported on another,[35–37] were handled as one RCT, consequently leaving a total of 14 trials that were included in the previous Cochrane review.

The 2004 update of the literature searches resulted in the identification of 25 references from MEDLINE and 35 from EMBASE. A total of 22 references were included in both MEDLINE and EMBASE, resulting in 38 references that met our inclusion criteria. After the selection and discussion step based on the titles, key words, and abstracts, both reviewers agreed that eight papers met the inclusion criteria.[24,27,28,42–46] Of these eight papers, two papers by Dalichau *et al*[42,43] reported on the same study, as did the papers by Lønn *et al*[44] and Glomsrød *et al*[45] and two papers by Indahl *et al*.[27,28] Papers that reported on the same study were handled as one RCT, leaving five eligible studies[24,27,28,42–46] that could be included in the updated review. Consequently, a total of 19 studies were included.

Six studies were excluded (Table 3) because they either did not meet our inclusion criteria for this review or had fatal errors in their design.[48–56] We excluded the studies by Stankovic and Johnell[48,49] that were included in the original Cochrane review. This study included a "Mini Back School," consisting of only one 45-minute session, with back care education as control group. The back school did not include exercises and therefore did not meet our definition of a back school. In the studies by Molde Hagen *et al*,[50,51] a modified version of a mini back school was applied to individuals instead of groups of patients. The studies by Frost *et al*[52,53] did not use an appropriate contrast for the back school. In the study by Morrison *et al*,[54] each group was assessed only once, the control group at the beginning of the program and the back school group at the end. In the study by Mucha and Winkler,[55] the study population was randomized into a back school and control group, but the differences between these groups were not analyzed. Instead, the authors only presented data for the back school group. In the feasibility study by Roques *et al*,[56] not all patients were randomly allocated to treatment; consequently, this study did not meet the criteria of an RCT.

Methodologic Quality

For the assessment of the methodologic quality of the trials, we combined the information from all papers reporting on the same trial. The methodologic quality of the 19 trials is presented in Table 4. Initially, there was

Table 5. Clinical Relevance

Study	1	2	3	4	5
Berquist[25]	+	+	+	+	?
Berwick[22]	+	−	−	−	?
Dalichau[42,43]	+	+	+	+	?
Donchin[40]	+	−	−	?	?
Herzog[47]	−	−	−	+	?
Hsieh[24]	+	+	+	−	+
Hurri[35–37]	−	−	+	−	?
Härkäpää[31–34]	−	−	+	−	?
Indahl[27,28]	−	+	−	−	?
Keijsers[38]	−	−	+	−	?
Keijsers[41]	−	+	+	−	?
Klaber Moffett[29]	+	+	+	−	?
Lankhorst[21]	−	−	+	−	?
Leclaire[23]	+	+	+	+	?
Lindequist[26]	+	−	−	−	?
Linton[39]	−	−	+	+	?
Lønn[44,45]	+	−	+	−	?
Penttinen[46]	+	−	+	−	?
Postacchini[30]	−	−	−	−	?

disagreement between the reviewers on 58 (28%) of the 209 items scored. Most disagreements were resolved in discussion. The third reviewer only had to make a final decision once. We sent the results of our quality assessment to the (first) authors of the RCTs, asking them if they agreed with our score, and, if not, to state the reasons. We also asked them for additional information if our final score was "unclear." Ten authors responded to our request.[21,24,26,27–29,31–37,39,40,44,45] The final scores, based on the comments and additional information of the authors of the studies are presented in Tables 1 and 2. One author agreed with our score and nine provided additional information. We changed 26 scores, 11 from unclear to positive, 9 from unclear to negative, and 6 from negative to positive. Only six studies had six or more positive scores.[23,24,27–29,39,44,45] The most prevalent methodologic shortcomings appeared to be Items 4 and 9, *i.e.*, in none of the RCTs were the patients or care providers blinded. Other methodologic flaws that occurred in more than half of the studies were: an inappropriate method of randomization (item 1a), no concealment of treatment allocation (item 1b), no measures taken in the study design to avoid co-interventions (item 3), no blinding of observers (item 5), and either no satisfactory compliance of interventions or no measurement of compliance at all (item 10).

Statistical Pooling

RCTs that provide sufficient and similar information on study setting, reference group, study population, and LBP characteristics are required for statistical pooling, as proposed in Methods. Of the 19 RCTs, five studies did not provide any usable information because outcome measures like spinal mobility or the Sickness Impact Profile were sometimes used or they lacked necessary data.[22,26,30,38,40] Furthermore, six of the remaining 14 RCTs could not be included in the statistical pooling because sensible subgroups could not be produced. Three of these studies included acute and subacute patients but used different reference groups. Three studies included patients with a mix of acute and chronic LBP. Consequently, eight RCTs that included chronic LBP patients with or without radiation were available for statistical pooling. With respect to the timing and presence of outcome measurements of these eight RCTs, we initially tried to pool the data for the subgroups described below.

1. Pain
   a) back school *versus* other treatments (exercises, spinal or joint manipulation, myofascial therapy, instructions or advice or another type of back school): for intermediate-term[29,31–37] and long-term follow-up.[31–37]
   b) back school *versus* WLC or placebo: for short-term[21,39,41–43] and intermediate-term follow-up.[21,39,41–43]
2. Functional status
   a) back school *versus* other treatments for short-term,[29,31–34] intermediate-term,[35–37,46] and long-term follow-up.[31–37,46]
   b) back school *versus* WLC or placebo for intermediate-term follow-up.[39,41]
3. Return to work
   a) back school *versus* WLC or placebo.[27,28,39,41–43]

Disappointingly, all studies either reported means without standard deviations or did not report group size. Because of this lack of information, we were unable to statistically pool the data and consequently performed a best evidence synthesis.

Effectiveness of Back Schools

1a) Back Schools *versus* Other Treatments for Acute/Subacute LBP. Some RCTs,[23,25–28] including some of high quality,[23,27,28] studied differences between a back school and other treatments for acute and subacute LBP patients. One high-quality RCT[27,28] reported positive intermediate- and long-term outcomes, and the other high and low quality studies reported no differences in short-, intermediate-, and long-term outcomes between those receiving back schools and other treatments.

There is conflicting evidence (4 trials; 1,418 patients) on the effectiveness of back schools compared with other treatments for acute and subacute LBP on pain, functional status, recovery, recurrences, and return to work (short-, intermediate-, and long-term follow-up).

1b) Back Schools *versus* Other Treatments for Chronic LBP. Six studies were identified that evaluated the effectiveness of back schools compared with other conservative treatments for chronic LBP.[29–37,40,46] Other conservative treatments were: exercises, spinal or joint manipulation, myofascial therapy, and some kind of instructions or advice. The high-quality study[28] and four low-quality studies[30,31–37,46] showed better short- and intermediate-term pain relief and improvement in functional status for the back school group. Three low-quality studies did not find any differences in long-term outcomes.[31–37,40] There is moderate evidence (5 trials; 1,095 patients) that a back school is more effective than other treatments for patients with chronic LBP for the outcomes pain and functional status (short- and intermediate-term follow-up). There is moderate evidence (3 trials; 822 patients) that there is no difference in long-term pain and functional status between those receiving back school and other treatments, for patients with chronic LBP.

2a) Back Schools *versus* WLC or Placebo Interventions for Acute/Subacute LBP. Only one RCT compared back school with placebo, *i.e.*, shortwaves at the lowest intensity, for patients with acute and subacute LBP and showed better short-term recovery and return to work for the back school treatment group.[25] No other short- or long-term differences were found.

Therefore, there is limited evidence (1 trial; 217 patients) that back school is more effective than shortwaves

at the lowest intensity for patients with acute and subacute LBP on recovery and return to work (short-term follow-up). There is limited evidence (1 trial; 217 patients) that there is no difference in short-term pain and long-term recurrences between the back school and shortwaves at the lowest intensity, for patients with acute and subacute LBP.

2b) Back Schools *versus* WLC or Placebo Interventions for Chronic LBP. Eight RCTs were identified for this subgroup analysis,[21,30,38–45] including two high-quality trials.[39,44,45] Seven RCTs reported a mix of positive results, with no differences in short- and intermediate-term outcomes.[21,30,38,39,41–45] One high-quality study found positive long-term outcomes on functional status and return to work[44,45] and two did not find any long-term differences.[21,40]

There is conflicting evidence (8 trials; 826 patients) on the effectiveness of back schools compared with WLC or placebo interventions on pain, functional status, and return to work (short-, intermediate-, and long-term follow-up), for patients with chronic LBP.

Back Schools in Occupational Settings. Nine studies (three high-quality studies[23,27–28,39] and six low quality studies,[25,31–37,40,42,43,46]) included patients from an occupational setting.

3a) Back Schools in Occupational Settings *versus* Other Treatments for Acute/Subacute LBP. Three studies,[23,25,27,28] including two high-quality studies,[23,27,28] examined the effect of a back school compared with other treatments for acute and subacute patients. One high-quality study found positive intermediate- and long-term results for the back school.[27,28] The other high- and low-quality RCTs found no short-, intermediate-, or long-term differences between the back school and other treatments.

There is conflicting evidence (3 trials; 1,362 patients) on the effectiveness of back schools compared with other treatments for acute and subacute LBP on return to work (short-, intermediate-, and long-term follow-up). There is moderate evidence (2 trials; 387 patients) that there is no difference in short-term pain for patients with acute and subacute LBP, between those who received the back school and other treatments. There is limited evidence that there is no difference in short-term functional status (1 trial; 170 patients), long-term recurrences (1 trial; 217 patients) and intermediate and long-term pain and functional status (1 trial; 170 patients) between the back school and other treatments, for patients with acute and subacute LBP.

3b) Back Schools in Occupational Settings *versus* Other Treatments for Chronic LBP. Four studies examined the effects of a back school compared with other treatments for chronic LBP patients.[31–37,40,46] One RCT studied short- and long-term differences,[31–34] two studies, intermediate- and long-term differences,[35–37,46] and one study, only long-term differences.

There is moderate evidence (3 trials; 764 patients) that a back school is more effective than other treatments for patients with chronic LBP for pain and functional status (short- and intermediate-term follow-up). There is conflicting evidence (4 trials; 906 patients) on the effectiveness of back schools compared with other treatments for chronic LBP on pain and functional status (long-term follow-up).

4a) Back Schools in Occupational Settings *versus* WLC or Placebo Interventions for Acute/Subacute LBP. Results for 4a are the same as the results presented under 2a because the trial was conducted in an occupational setting.

4b) Back Schools in Occupational Settings *versus* WLC or Placebo Interventions for Chronic LBP. Three RCTs examined the effect of a back school compared with WLC for chronic LBP.[39,40,42,43] Two studies found positive short- and intermediate-term results,[39,42,43] and one did not find any long-term differences.[40]

There is moderate evidence (2 trials; 186 patients) that a back school is more effective than WLC for patients with chronic LBP for pain and return to work (short- and intermediate-term follow-up). There is limited evidence (1 trial; 142 patients) that there is no difference in long-term incidence of LBP episodes between back school and WLC for patients with chronic LBP.

Sensitivity Analysis

We defined high-quality as meeting six of the 11 internal validity criteria. A sensitivity analysis was carried out using different cutoff points, *i.e.*, high quality defined as either five or seven of the 11 items scored positive. Only one study met seven or more of the criteria out of 11,[29] and nine studies met at least five of the 11 criteria.[23–25,27–29,31–34,39,40,44,45] If high quality was defined as seven or more items had to meet our criteria, the strength of the evidence would remain the same. If high quality was defined as five or more items meeting our criteria, there would be strong evidence that back schools were more effective than other treatments for chronic back pain at short-term follow-up and strong evidence that there was no difference in long-term follow-up for the same comparison. Furthermore, there would be strong evidence for the effectiveness of back schools compared with WLC or placebo for patients with chronic back pain. Changing the high-quality level to meeting five or more items would not contaminate the conclusions with respect to the strength of evidence for back schools in an occupational setting. If we related high quality to an adequate concealment of treatment allocation, which could be identified in three RCTs,[29,39,40] the strength of the evidence did not change.

Clinical Relevance

The clinical relevance of the studies was assessed independently by two reviewers, by scoring the five questions presented in Methods. See Table 5 for the scores. Disagreement between the two reviewers existed on 32 (34%) of the 95 clinical relevance scores. Overall, none of the clinical relevance scores of the RCTs was sufficient.

None of the RCTs scored a positive on all clinical relevance items, and only four RCTs[23–25,42,43] scored positively on four of the five questions. The most negative scores were assigned to the questions about the description of the interventions and treatment settings and the clinical importance of the effect size. Furthermore, the majority of studies scored unclear for the question regarding the likely treatment benefits, indicating a clear lack of information on this topic.

■ Discussion

In this review update, 19 RCTs were included that evaluated the effectiveness of back schools for nonspecific LBP. Positive short-term effects of back schools were seen in most of the RCTs. Positive long-term effects of back schools were only reported in four RCTs.[25,27,28,44–46] Of particular note was the heterogeneity among studies with respect to study populations, content of back schools, type of control interventions, and outcome measurements. Studies also differed in cultural setting, with most studies conducted in Scandinavia. This means that, with respect to the generalizability of the results, cultural differences in healthcare and social security systems have to be considered. We qualitatively assessed the strength of evidence of the included RCTs by applying a methodologic criteria list and best evidence synthesis.[16] Because of the divergence in study designs and the generally low methodologic quality scores of the RCTs, it was not possible to perform statistical pooling of the data. According to the best evidence synthesis, we could not identify strong evidence for any type of back school treatment. We concluded that there is moderate evidence that back schools have better short- and intermediate-term effects than other treatments for recurrent and chronic LBP for pain and functional status. Furthermore, there is moderate evidence that back schools in an occupational setting are more effective than other treatments, placebo, or WLC for chronic LBP for pain, functional status, and return to work during short- and intermediate-term follow-up. The clinical relevance of the RCTs, scored by answering five questions, was insufficient.

The generally low methodologic quality scores of the included RCTs, originating from the many shortcomings in trial design and performance, was striking. After weighing the methodologic quality of the studies, six RCTs could be identified as being of high methodologic quality,[23,24,27–29,39,44,45] leaving 13 RCTs of low methodologic quality. Compared with the cutoff value of meeting at least six criteria to be considered high quality, the mean total quality score of the 19 included RCTs was 3.8, which is a low score. The most commonly identified methodologic deficiencies were the lack of blinding of patients (scored negative in all 19 RCTs), observers, and care providers (scored negative in all 19 RCTs), an inappropriate method of randomization (scored negative or unclear in 13 RCTs), inadequate concealment of treatment allocation (scored negative or unclear in 16 RCTs), lack of avoidance of co-interventions (scored negative or unclear in 16 RCTs), and unsatisfactory compliance with the interventions (scored negative or unclear in 16 RCTs). Empirical evidence reports an existing association between inadequate concealment of treatment allocation and lack of double-blinding (blinding of patients and observers) with bias.[57–59] In only 6 of 19 RCTs, a clear description of the randomization procedure was available.[23–25,29,39,44,45] More disappointing was that in only 3 of 19 studies was there a clear description of the treatment allocation.[29,39,40] The reported methodologic limitations are not unique for clinical trials evaluating the efficacy of back schools but have also been demonstrated in trials on other conservative treatments for LBP. However, reports of RCTs should be accurate and complete so that readers can evaluate the internal and external validity of the trial. In this review, we changed 24 quality assessment scores after additional information was provided by eight authors, which indicates that the quality of the report is not similar to the quality of the trial. Contacting the authors for additional information is only one solution to this problem and has the disadvantage that it may be difficult to contact authors of trials that have been published for several years. The quality of future RCTs in the field of back pain should be improved to reduce bias in systematic reviews, as it has been demonstrated that statistical pooling of low-quality trials results in overestimation of treatment effects.

For clinicians and other caregivers, it is essential not only to be informed about the effectiveness of back schools, but also about the characteristics of the included patients, clinical relevance of the effect size, and content of the programs, to determine the clinical relevance of the studies for their patient population. In this review, we assessed the clinical relevance of the RCTs by scoring five questions recently recommended by the Cochrane Back Review Group[16] and published earlier by Shekelle *et al*[60] and Guyatt *et al*.[61] The majority of RCTs did not score sufficiently on the questions, especially Questions 2 and 4, which described the intervention and treatment settings and the clinically relevant effect size. RCTs either reported briefly about the content of the intervention or failed to report essential information about the type, intensity, or performance of the exercises. Also apparent were the widely variations in the content and components of the interventions. This may explain the differences in interpretation of the items between the two reviewers, reflected by the disagreement score of 34%. It is recommended that future RCTs explicitly describe clinically relevant aspects of intervention programs, alongside the development of stricter criteria to validly judge the clinical relevance. This information is important to eventually identify which element of a back school program is responsible for changes in outcome for a specific type of LBP patient, to improve clinical care.[63,64]

As the effect of back schools is likely to be small, a meta-analysis in which information from multiple RCTs is combined could provide reliable evidence about the

effectiveness of back schools. In this systematic review, we tried to perform a formal meta-analysis of the data studied. However, most of the studies reported insufficient information about means, standard deviations, or group size, with the result that a quantitative summary of the data was impossible. Furthermore, it is not yet commonly accepted for RCTs in LBP research to measure the four recommended outcome measures of pain, functional status, improvement, and return to work. The majority of included studies in this review reported information on pain, namely, 16 of the 19 included RCTs. However, a limited number of studies reported on functional status and return to work; only 7 and 4 of the 19 studies, respectively. Recently, several attempts were made to recommend and standardize the use of these outcome measures for LBP research.[17,18] Hopefully, this will enhance the comparability of future RCTs, which will result in more sensible subgroup analyses.

Several biases can be introduced by the literature search and selection procedure. We might have missed relevant unpublished trials, which are more likely to be small studies with nonsignificant or negative results, leading to publication bias.[64] Screening references of identified trials and systematic reviews may result in an overrepresentation of positive studies in the review because trials with a positive result are more likely to be referred to in other publications, leading to reference bias.[65] We tried to identify RCTs published in English, Dutch, French, and German, but English trials most often met our inclusion criteria and were included in the systematic review. Two RCTs of back schools were identified that were published in German, one low-quality study[42,43] and another excluded study, because the analyses were only presented for the back school group.[55] It has been demonstrated in fields other than back pain, that among published studies, those with significant results are more likely to get published in English, leading to language bias.[66] Because one of the main principles of a systematic review is to include all available evidence, identification of all trials is important to the validity of a systematic review. Biases such as publication, reference, or language bias can be avoided by the use of prospective registries of trials. However, the final decision to include, for example, unpublished trials or trials published in languages other than English, may be based on practical reasons rather than methodologic ones.

At present, cost-effectiveness analyses of back schools have not been conducted alongside RCTs. Two nonrandomized trials on cost-effectiveness did not show any significant differences between back schools and no treatment.[67,68] The length of back schools included in this review varied from a Swedish back school containing four 45-minute sessions to more intensive back schools with a 3- to 5-week stay in a specialized center. These more intensive back schools are likely to be more expensive. If back schools of different intensity are similarly effective, the next step should be to gain insight into the cost-effectiveness of these different types of back schools. Correctly, Goossens and Evers[69] concluded, in their publication of a review of economic evaluations of all kinds of interventions for back pain, that there is a need for improvement of the methodologic quality of the cost-effectiveness studies.

Review articles offer clinicians and health policy makers the opportunity to cope with the exponentially increasing number of medical publications like RCTs. Conclusions originating from systematic reviews contribute largely to the development and implementation of practical LBP guidelines to enhance clinical care. For a long time, clinicians have been provided with evidence from nonsystematic narrative reviews. The current interest in evidence-based medicine has led to an extensive increase in the publication of systematic summaries of RCTs. Randomized controlled trials are generally considered to be the paradigm of intervention research, that is, the strongest scientific proof of the effectiveness of an intervention. Recently, recommendations for the reporting of RCTs, the CONSORT statement, were published, which have been adopted by several leading medical journals and included in their instructions to authors.[70,71] Although systematic reviews of RCTs have their limitations, there seems to be consensus that there is a need for them and that they need to be conducted as carefully as the studies they report.[72,73] The recent development and publication of methodical guidelines for systematic reviews of RCTs in the field of LBP offer guidance to researchers preparing, conducting, or reporting a systematic review and to readers evaluating these reviews.[16] It is still not possible to give a definite answer on the question on whether back schools are an effective treatment for (subgroups of) LBP patients. Prospective improvement in the quality of reporting of RCTs seems to be the best option to reduce bias in future systematic reviews and to lead to strong levels of evidence.

Implications for Practice

There is moderate evidence that back schools conducted in occupational settings seem to be more effective for patients with recurrent and chronic LBP (as opposed to patients from the general population or primary/secondary care) than other treatments, placebo, or WLC for pain, functional status, and return to work during short- and intermediate-term follow-up. The most promising interventions consisted of a modification of the Swedish back school and were quite intensive (a 3- to 5-week stay in a specialized center).

■ Conclusion

We identified 19 RCTs (3,584 patients) that evaluated the effectiveness of back schools. Most of the studies included in this review showed methodologic deficiencies. Clearly, there is a need for future high-quality RCTs to determine which type of back school is the most effective for LBP patients. Furthermore, future RCTs should include an evaluation of the cost-effectiveness of back

schools and consider the clinical relevance of the trial more during study design and performance.

■ Key Points

- A systematic review of 19 randomized controlled trials that evaluated the effectiveness of back schools was performed.
- Most of the randomized controlled trials were of low methodologic quality and did not score sufficiently on their clinical relevance.
- There is moderate evidence that back schools conducted in occupational settings seem to be more effective for patients with recurrent and chronic low back pain (as opposed to patients from the general population or primary/secondary care) than other treatments, placebo, or waiting list controls for pain, functional status, and return to work during short- and intermediate-term follow-up.
- Future randomized controlled trials should include an evaluation of the cost-effectiveness of back schools and consider the clinical relevance of the trial during study design and performance.

 *tables*

Acknowledgments

The authors thank prof. Dr. L. M. Bouter for his insightful suggestions for the original review and Vicki Pennick, Back Group Coordinator, for her assistance and time in reviewing the manuscript.

References

1. Andersson GBJ. Epidemiological features of chronic low-back pain. *Lancet* 1999;354:581–5.
2. Frymoyer JW. Quality: an international challenge to the diagnosis and treatment of disorders of the lumbar spine. *Spine* 1993;18:2147–52.
3. Maniadakis N, Gray A. The economic burden of back pain in the UK. *Pain* 2000;84:95–103.
4. van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in the Netherlands. *Pain* 1995;62:233–240.
5. Waddell G. A new clinical model for the treatment of low back pain. *Spine* 1987;12:632–644.
6. Spitzer WO, LeBlanc FE, Dupuis M. Scientific approach to the assessment and management of activity-related spinal disorders. *Spine* 1987;7(suppl): 1–59.
7. van Tulder MW, Assendelft WJJ, Koes BW, et al. Method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group for Spinal Disorders. *Spine* 1997;22:2323–30.
8. Koes BW, van Tulder MW, Ostelo R, et al. Clinical guidelines for the management of low back pain in primary care: an international comparison. *Spine* 2001;26:2504–13.
9. Waddell G, Burton AK. Occupational health guidelines for the management of low back pain at work: evidence review. *Occup Med* (*Lond*) 2001;51: 124–35.
10. Staal JB, Hlobil H, van Tulder MW, et al. Occupational health guidelines for the management of low back pain: an international comparison. *Occup Environ Med* 2003;60:618–26.
11. Zachrisson-Forssell M. The Swedish back school. *Physiotherapy* 1980;66: 112–4.
12. Zachrisson-Forsell M. The back school. *Spine* 1981;6:104–6.
13. van Tulder MW, Esmail R, Bombardier C, et al. Back schools for nonspecific low back pain (Cochrane Review). In: *The Cochrane Database Library*, Issue 3, 2003. Oxford: Update Software.
14. Cohen JE, Goel V, Frank JW, et al. Group education interventions for people with low back pain. *Spine* 1994;19:1214–22.
15. Koes BW, van Tulder MW, van der Windt DAWM, et al. The efficacy of back schools: a review of randomised clinical trials. *J Clin Epidemiol* 1994;47: 851–62.
16. van Tulder M, Furlan A, Bombardier C, et al. Updated method guidelines for systematic reviews in the Cochrane Collaboration Back Review Group. *Spine* 2003;28:1290–9.
17. Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. *Spine* 2000;25(suppl): 3100–3.
18. Deyo RA, Battie M, Beurskens AJ, et al. Outcome measures for low back pain research: a proposal for standardized use. *Spine* 1998;23:2003–13.
19. Deyo RA. Measuring the functional status of patients with low back pain. *Arch Phys Med Rehabil* 1988;69:1044–53.
20. Robinson KA, Dickersin K. Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. *Int J Epidemiol* 2002;31:150–3.
21. Lankhorst GJ, van der Stadt RJ, Vogelaar TW, et al. The effect of the Swedish back school in chronic idiopathic low-back pain. *Scand J Rehabil Med* 1983; 15:141–5.
22. Berwick DM, Budman S, Feldstein M. No clinical effect of back schools in an HMO: a randomized prospective trial. *Spine* 1989;14:339–44.
23. Leclaire R, Esdaile JM, Suissa S, et al. Back school in a first episode of compensated acute low back pain: a clinical trial to assess efficacy and prevent relapse. *Arch Phys Med Rehabil* 1996;77:673–9.
24. Hsieh CY, Adams AH, Tobis J, et al. Effectiveness of four conservative treatments for subacute low back pain: a randomized clinical trial. *Spine* 2002;27:1142–8.
25. Bergquist-Ullman M, Larsson U. Acute low-back pain in industry. *Acta Orthop Scand* 1977;170(suppl):1–117.
26. Lindequist SL, Lundberg B, Wikmark R, et al. Information and regime at low-back pain. *Scand J Rehabil Med* 1984;16:113–6.
27. Indahl A, Velund L, Reikeraas O. Good prognosis for low back pain when left untampered: a randomized clinical trial. *Spine* 1995;20:473–7.
28. Indahl A, Haldorsen EH, Holm S, et al. Five-year follow-up study of a controlled clinical trial using light mobilization and an informative approach to low back pain. *Spine* 1998;23:2625–30.
29. Klaber Moffett JA, Chase SM, Portek I, et al. 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.
30. Postacchini F, Facchini M, Palieri P. Efficacy of various forms of conservative treatment in low-back pain: a comparative study. *Neuro-Orthopedics* 1988; 6:28–35.
31. Härkäpää K, Järvikoski A, Mellin G, et al. A controlled study on the outcome of inpatient and outpatient treatment of low-back pain. Part I. *Scand J Rehabil Med* 1989;21:81–9.
32. Härkäpää K, Mellin G, Järvikoski A, et al. A controlled study on the outcome of inpatient and outpatient treatment of low-back pain. Part III. *Scand J Rehabil Med* 1990;22:181–8.
33. Mellin G, Hurri H, Härkäpää K, et al. A controlled study on the outcome of inpatient and outpatient treatment of low back pain. Part II. *Scand J Rehabil Med* 1989;21:91–5.
34. Mellin G, Härkäpää K, Hurri H, et al. A controlled study on the outcome of inpatient and outpatient treatment of low back pain. Part IV. *Scand J Rehabil Med* 1990;22:189–94.
35. Hurri H. The Swedish back school in chronic low-back pain. Part I. Benefits. *Scand J Rehabil Med* 1989;21:33–40.
36. Hurri H. The Swedish back school in chronic low-back pain. Part II. Factors predicting the outcome. *Scand J Rehabil Med* 1989;21:41–4.
37. Julkunen J, Hurri H, Kankainen J. Psychological factors in the treatment of chronic low back pain. *Psychother Psychosom* 1988;50:173–81.
38. Keijsers JFEM, Groenman NH, Gerards FM, et al. A back school in the Netherlands: evaluating the results. *Patient Education Counseling* 1989;14: 31–44.
39. Linton SJ, Bradley LA, Jensen I, et al. The secondary prevention of low back pain: a controlled study with follow-up. *Pain* 1989;36:197–207.
40. Donchin M, Woolf O, Kaplan L, et al. Secondary prevention of low-back pain: a clinical trial. *Spine* 1990;15:1317–20.
41. Keijsers JFME, Steenbakkers WHL, Meertens RM, et al. The efficacy of the back school: a randomized trial. *Arthritis Care Res* 1990;3:204–9.
42. Dalichau S, Perrey RM, Solbach T, et al. Erfahrungen bei der Durchführung eines berufsbezogenen Rückenschulmodells im Baugewerbe. *Zentralbl Arbeitsmedizin* 1998;48:72–80.
43. Dalichau S, Scheele K, Perrey RM, et al. Ultraschallgestützte Haltungs- und Bewegungsanalyse der Lendenwirbelsäule zum Nachweis der Wirksamkeit einer Rückenschule. *Zentralbl Arbeitsmedizin* 1999;49:148–56.

44. Lønn JH, Glomsrød B, Soukup MG, et al. Active back school. Prophylactic management for low back pain: a randomized controlled 1-year follow-up study. *Spine* 1999;24:865–71.
45. Glomsrød B, Lønn JH, Soukup MG, et al. Active back school, prophylactic management for low back pain: three-year follow-up of a randomized controlled trial. *J Rehabil Med* 2001;33:26–30.
46. Penttinen J, Nevala-Puranen N, Airaksinen O, et al. Randomized controlled trial of back school with and without peer support. *J Occup Rehabil* 2002; 12:21–9.
47. Herzog W, Conway PJW, Willcox BJ. Effects of different treatment modalities on gait symmetry and clinical measures for sacroiliac joint patients. *J Manipulative Physiol Ther* 1991;14:104–9.
48. Stankovic R, Johnell O. Conservative treatment of acute low-back pain. A prospective randomized trial: McKenzie method of treatment versus patient education in 'mini-back school.' *Spine* 1990;15:120–3.
49. Stankovic R, Johnell O. Conservative treatment of acute low back pain: a 5-year follow-up study of two methods of treatment. *Spine* 1995;20:469–72.
50. 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;25:1973–6.
51. Molde Hagen E, 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;28:2309–15.
52. Frost H, Klaber Moffett JA, Moser JS, et al. Randomised controlled trial for evaluation of fitness programme for patients with chronic low back pain. *BMJ* 1995;310:151–4.
53. Frost H, Lamb SE, Klaber Moffett JA, et al. A fitness programme for patients with chronic low back pain: 2-year follow-up of a randomised controlled trial. *Pain* 1998;75:273–9.
54. Morrison GEC, Chase W, Young V, et al. Back pain: treatment and prevention in a community hospital. *Arch Phys Med Rehabil* 1988;69:605–9.
55. Mucha C, Winkler J. Back school for nursing staff. *Phys Rehab Kur Med* 1996;6:48–52.
56. Roques CF, Felez A, Marque P, et al. Feasibility of a back school assessment programme. *Ann Readapt Med Phys* 2002;45:257–64.
57. Chalmers TC, Celano P, Sacks HS, et al. Bias in treatment assignment in controlled clinical trials. *N Engl J Med* 1983;309:1358–61.
58. Colditz GA, Miller JN, Mosteller F. How study design affects outcomes in comparisons of therapy. I: Medical. *Stat Med* 1989;8:441–54.
59. Schultz KF, Chalmers I, Hayes RJ, et al. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408–12.
60. Shekelle PG, Andersson G, Bombardier C, et al. A brief introduction to the critical reading of the clinical literature. *Spine* 1994;19(suppl):28–31.
61. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature: II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. *JAMA* 1994;271:59–63.
62. Mannion AF, Taimela S, Muntener M, et al. Active therapy for chronic low back pain. Part 1. Effects on back muscle activation, fatigability, and strength. *Spine* 2001;26:897–908.
63. Staal JB, Hlobil H, van Tulder MW, et al. Return-to-work interventions for low back pain: a descriptive review of contents and concepts of working mechanisms. *Sports Med* 2002;32:251–67.
64. Egger M, Davey Smith G. Bias in location and selection of studies. *BMJ* 1998;316:61–6.
65. Goetzsche PC. Reference bias in reports of drug trials. *BMJ* 1987;295: 654–6.
66. Egger M, Zellweger-Zähner T, Schneider M, et al. Language bias in randomised controlled trials published in English and German. *Lancet* 1997;350: 326–9.
67. Brown KC, Sirles AT, Hilyer JC, et al. Cost-effectiveness of a back school intervention for municipal employees. *Spine* 1992;17:1224–8.
68. Versloot JM, Rozeman A, van Son AM, et al. The cost-effectiveness of a back school program in industry: a longitudinal controlled field study. *Spine* 1992; 17:22–7.
69. Goossens MEJB, Evers SMAA. Economic evaluation of back pain interventions. *J Occup Rehabil* 1997;7:15–32.
70. Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *JAMA* 1996;276: 637–9.
71. Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001;357:1191–4.
72. Moher D, Olkin I. Meta-analysis of randomized controlled trials: a concern for standards. *JAMA* 1995;274:1962–4.
73. Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomised controlled trials. *Lancet* 1998;351:47–52.