

Prediction of Sickness Absence in Patients with Chronic Low Back Pain: A Systematic Review

Wietske Kuijer · Johan W. Groothoff · Sandra Brouwer ·
Jan H. B. Geertzen · Pieter U. Dijkstra

Published online: 17 June 2006
© Springer Science+Business Media, Inc. 2006

Objectives: To provide evidence of predictors for sickness absence in patients with non-specific chronic low back pain (CLBP), distinguishing predictors aimed at the decision to report sick (absence threshold) and decision to return to work (return to work threshold). **Methods:** Medical and psychological databases were searched, as well as citations from relevant reviews. In- and exclusion criteria were applied. Two reviewers assessed the methodological quality of the papers independently. **Results:** Many different predictors were studied, and few factors were studied more than once. Consistent evidence was found for own expectations of recovery only as predictor for the decision to return to work. Patients with higher expectations had less sickness absence at the moment of follow-up measurement. As expected, different predictors were found aiming at the absence threshold or the return to work threshold. Furthermore, predictors varied also with the measurement instruments used, timing of follow-up measurements, and definition of outcomes. Until now, too few studies are available to overcome several potential sources of heterogeneity. **Conclusions:** No core set of predictors exists for sickness absence in general. The characteristics of the study including the decision to report sick or to return to work determined the influence of several predictors on sickness absence in patients with CLBP. Further research

W. Kuijer · J. W. Groothoff · S. Brouwer · J. H. B. Geertzen · P. U. Dijkstra
Center for Rehabilitation, University Medical Center Groningen, University of Groningen,
Groningen, The Netherlands

W. Kuijer · J. W. Groothoff · S. Brouwer · J. H. B. Geertzen · P. U. Dijkstra
Northern Center for Healthcare Research, University Medical Center Groningen, University of Groningen,
Groningen, The Netherlands

P. U. Dijkstra
Department of Oral and Maxillofacial Surgery, University Medical Center Groningen,
University of Groningen,
Groningen, The Netherlands

W. Kuijer (✉)
Han University, Faculty of Health and Social Studies, P.O. Box 6960,
NL-6503 GL Nijmegen, The Netherlands
e-mail: wietske.kuijer@han.nl

and use of a core set of measurements and uniform definitions are needed to predict sickness absence and return to work in patients with CLBP.

Keywords Chronic low back pain · Prediction · Sick leave · Systematic review

Introduction

Many people suffer from low back pain (LBP) once in their life [1]. In 80 to 95% of patients with LBP, no specific origin of the back pain can be found [2]. This LBP is called non-specific, simple or mechanical back pain. Non-specific LBP often develops spontaneously, and mostly resolves within 4–6 weeks after onset [3, 4]. In some cases, the back pain persists and the pain becomes chronic. This transition from acute to chronic LBP (CLBP) is complex. Many individual, psychosocial and work related factors, the so called ‘yellow flags’ (e.g. fear avoidance behavior, catastrophising, passive attitude to rehabilitation, depression, anxiety, psychosocial aspects of work, compensation) may contribute to the persistence of LBP [5]. Patients with CLBP account for 75 to 90% of the socio-economic costs of LBP [6], mainly a consequence of healthcare interventions and work incapacity associated with CLBP [7]. To lower these costs, absence from work associated with CLBP should be reduced.

Not all patients with CLBP are absent from work. All workers have a certain reluctance before reporting sick, the so-called absence threshold. The absence threshold is determined by the opportunity for absence or work (such as sanctions, financial impact) and the need for absence or work (for instance severity of LBP, disability, job demands, job satisfaction) [8]. In addition, in the recovering stage, workers also have to decide to return to work (RTW), and have to overcome the RTW threshold [8]. This RTW threshold is more than the reflection image of the absence threshold. Additional factors that play a role during the absence period may influence the decision of the worker to RTW, such as recommendation of a company doctor, influence of healthcare and work factors such as the ability to work less hours. This indicates that the absence threshold and the RTW threshold can be influenced by different factors (Fig. 1). Therefore, it is important to identify predictors for sickness absence aimed at the absence threshold and the RTW threshold.

Several studies on predictive factors for sickness absence have been conducted, however, many studies were aimed on predictive factors in the acute stage of LBP. Due to the presence of different individual beliefs and behaviors in patients with CLBP (‘the yellow flags’), it is likely, that predictors for absence are different in chronic patients. Only a few systematic reviews have been conducted recently on predictors of sickness absence in patients with CLBP. These reviews included only studies aimed at intervention strategies, leaving out

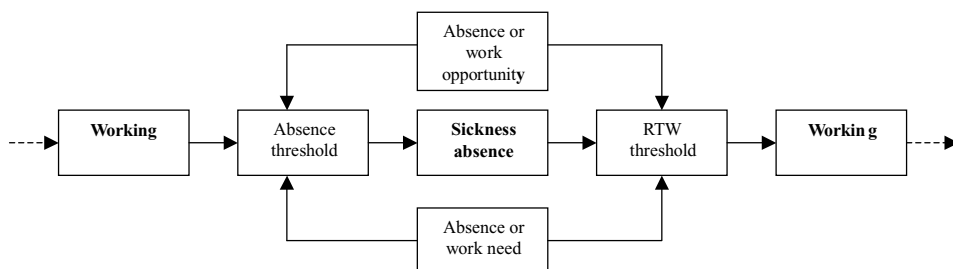


Fig. 1 Model for sickness absence. Based on: Allegro and Veerman [8]

prognostic cohort studies without intervention [9–11], limited the inclusion of CLBP patients on maximum duration of LBP [12, 13], did not differentiate between acute, sub acute or chronic patients [14], or the literature search was not described clearly [12]. In addition, none of the studies made the distinction between the absence threshold and the RTW threshold. Due to these limitations of previous reviews and the recent growing number of prospective research on this topic, a systematic review on risk factors for sickness absence in patients with CLBP was desirable distinguishing predictors aimed at the absence threshold and RTW threshold.

The aim of this review was to provide an overview of predictors for sickness absence in patients with CLBP, for both the absence threshold and RTW threshold, by reviewing the literature systematically and assessing the methodological quality of the papers.

Materials and methods

Search strategy

The databases Medline, Embase, Cinahl, Amed, Psychinfo and Cochrane were searched from January 1980 (or first administered year from 1980) to October 2004. The search term “Low Back Pain” was entered as MesH term and as free text word. This term was combined with several MesH terms as “sick leave”, “absenteeism”, or “vocational rehabilitation”, and with several free text words as “return to work”, “job resumption” or “job loss”. A full description of the literature search and search items is presented in Appendix 1. Systematic reviews retrieved from the literature search on predictors of outcome were screened for additional relevant papers on sickness absence.

Inclusion and exclusion criteria

Application of in- and exclusion criteria were pilot tested by two reviewers (WK, PUD) and adjusted until consensus was reached. The final in- and exclusion criteria are presented in Appendix 2. One reviewer performed the first screening on the abstracts of the papers by (WK). Second, in- and exclusion criteria were applied to the full text of the papers (WK). Only papers written in Dutch, English or German were included for review. The other reviewer screened the included papers for methodological quality assessment on in- and exclusion criteria (PUD). Papers were not blinded for authors, institution, journal, results or conclusions for a practical point of view.

Criteria for methodological quality

Before quality assessment, application of criteria was pilot tested and adjusted until consensus was reached. All papers were assessed by two reviewers (WK, PUD), according to a methodological quality list for assessing prognostic studies, based on criteria used by the Cochrane Collaboration for observational studies [15], Borghouts [16] and Scholten Peeters [17] (Appendix 3).¹ In addition, Cochrane criteria for methodological quality assessment were used for assessing RCTs (Appendix 4)¹ [18]. Each criterion was graded as yes or no. If insufficient information was provided in the papers, the criterion was also assessed as no. If a paper referred to other sources for information, these sources were used to assess that specific criterion. Disagreement was discussed in a consensus meeting. When no consensus could be reached, a third reviewer (JHBG) was asked for a binding verdict.

¹ Operational definitions of both criteria lists are available from the authors upon request.

Data extraction

Papers that had an adequate description of in- and exclusion criteria, study population, predictive factors, had a follow-up duration ≥ 12 months, had an acceptable number of dropouts and had defined the outcome adequately were eligible for detailed review.¹ These predetermined criteria for prognostic studies (B, C, D, E, F and H, see Appendix 3) were chosen to include clinically homogeneous studies, which enables statistical pooling for analyses on predictive factors for sickness absence and RTW [19, 20]. If less than 5 papers fulfilled all 6 criteria, a quality score for prognostic studies was calculated by summing the ‘yes’ answers. Only high quality papers were included for review, i.e. papers with a quality score of prognostic studies ≥ 6 (maximum quality score = 9). These included papers were analyzed qualitatively, aimed at generating different levels of evidence for the predictors of outcome. Evidence generated from studies aimed at the absence threshold and the RTW threshold will be distinguished because differences might exist with respect to prognoses. Predictors were classified as demographic, life style, medical (history), pain, observed disability, self-reported disability, health beliefs, physical work demands, psychological work demands, emotions, expectations and interventions.

Levels of evidence

Four levels of evidence will be described for prognostic cohort studies or prognostic factors in RCT studies other than interventions; 1) consistent evidence, 2) limited evidence, 3) conflicting evidence and 4) no evidence. The overall conclusion of consistent evidence is defined as ‘two or more studies reporting associations with sickness absence, or at least 75% of the studies reporting similar conclusions.’ Limited evidence is present when only 1 study is available and it reports associations with sickness absence. Conflicting evidence is reported when $<75\%$ of available studies reported similar findings [21], or contradictory findings are present within one study. In case of conflicting findings in multiple studies, the available univariate analyses were disregarded in drawing the overall conclusion. No evidence is found if no associations with sickness absence are present in either one or in multiple studies. If one study used both univariate and multivariate analyses for the same predictor, the available univariate analyses were disregarded in generating the evidence. Results are statistically significant at $p \leq 0.05$. For interventions as predictors (in RCTs), levels of evidence will be described based on the U.S. Clinical Practice Guideline for Acute Low Back Pain in Adults [22]. Quality of RCTs was assessed summing the ‘yes’ answers. An RCT was considered to be of high quality if the methodological quality score for RCTs was more than 50% of the maximum quality score (RCT score of 6 or more, maximum of 11) [22]. Strong evidence is present when consistent findings in multiple high quality studies are reported. Moderate evidence is present when consistent findings in 1 high and 1 or more low quality studies are reported, or in multiple low quality studies. Limited evidence is present when only 1 study is available. Conflicting evidence is found when contradictory findings are reported in multiple studies, or contradictory findings within one study. In the case of multiple high quality studies, the available low quality studies were disregarded in drawing the overall conclusion.

Results

Selection of studies

The flow chart of the review process is shown in Fig. 2. In the first screening, 2137 abstracts were screened on in- and exclusion criteria. Fifty-five papers were included for methodological

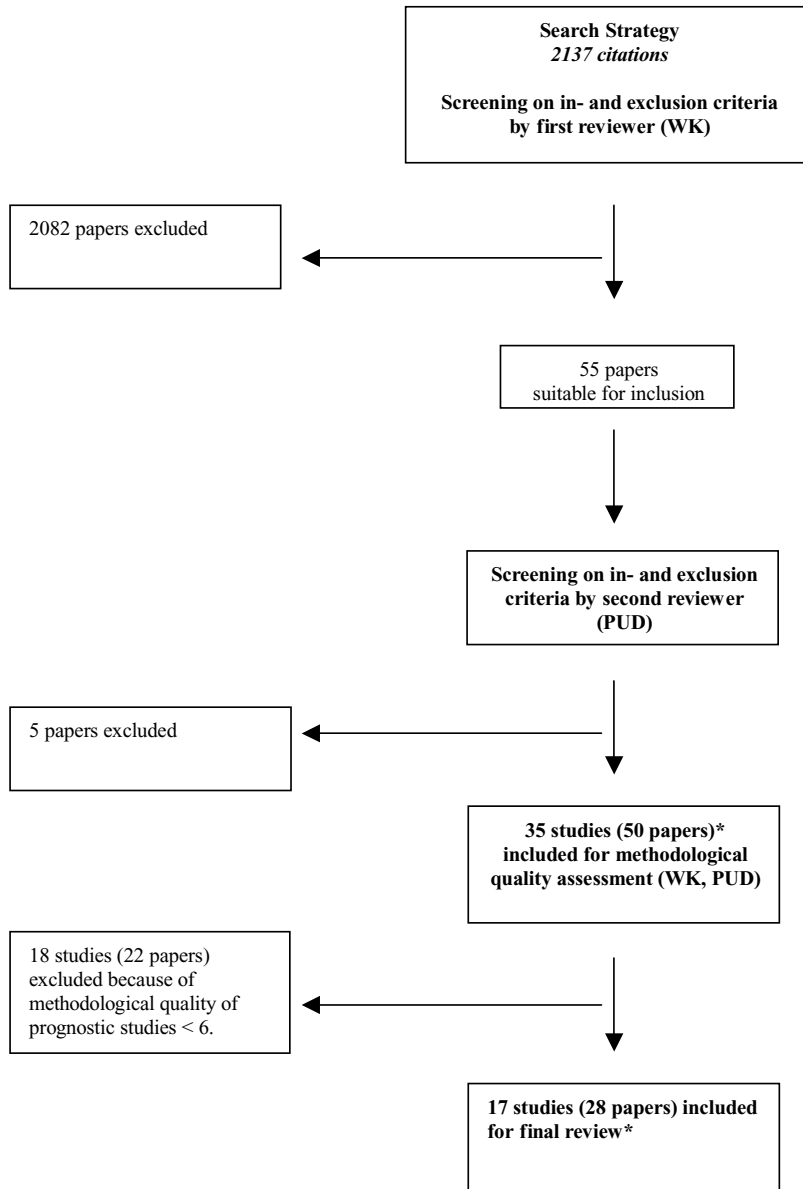


Fig. 2 Flowchart of the review process
 *Papers describing the same study cohort were considered as one study

quality assessment by the first reviewer. The second reviewer screened these 55 papers on in- and exclusion criteria by, and excluded another 5 papers.

Of the 50 papers included [23–72], several papers described the same cohort and also referred to the other papers for detailed information. Therefore, these papers were assessed on methodological quality simultaneously as one study, leaving 35 studies included for quality assessment. Overall absolute agreement of quality assessment was 84%, kappa (κ) = 0.67. The

absolute agreement of RCT quality assessment was 82% and $\kappa = 0.63$, and of prognostic cohort studies respectively 85% absolute agreement and $\kappa = 0.69$. Final consensus was reached without needing to consult the third reviewer. Of the 35 studies, only 3 studies had an adequate description of in- and exclusion criteria, study population, predictive factors, had a follow-up duration ≥ 12 months, had an acceptable number of dropouts and had defined the outcome adequately (predetermined criteria B, C, D, E, F and H, see Appendix 3).¹ [23, 33, 54, 66, 67] Therefore, studies that reached a prognostic quality score ≥ 6 were analyzed (Appendix 3). Eighteen studies did not reach the quality score ≥ 6 , leaving 17 studies for inclusion in the detailed review [23, 29–31, 33, 34, 43, 46–49, 51–54, 56–59, 63–69, 71, 72]. Methodological quality for the included studies on individual items is presented in Table 1 for both prognostic cohort and RCT quality assessment.

Only 7 studies described predictors other than intervention strategies (Table 3) [29, 30, 46–49, 51–53, 57–59, 63–65, 72]. Many different predictors were studied. Self-reported disability in activities of daily living (ADL) was studied most frequently (5 times), followed by previous duration of sick leave (4 times). Factors of life style, health beliefs and psychological work demands were studied only once mostly. The heterogeneity of the populations, predictive factors, intervention strategies and follow-up duration prevented us from statistical pooling of the study results. In addition, within and between studies, outcomes were presented as sickness absence at the moment of follow-up measurement (dichotomized) or as total number of days on sick leave during the follow-up period (Table 2). Therefore, it was decided to describe the predictors for sickness absence at the moment of follow-up measurement (dichotomous) and for total number of sick leave days in the follow-up period both for the absence threshold and the RTW threshold.

Fourteen studies described intervention strategies as predictor for sickness absence (Table 4) [23, 29, 31, 33, 34, 43, 46–49, 51–54, 56–59, 66–69, 71, 72]. In 2 studies, no difference between the absence threshold and RTW threshold could be made, due to studying a mixed population of patients at work and patients already sick listed [31, 43]. Of these groups, it was unknown which patients remained sick listed, which patients recovered and which patients deteriorated. One study did not apply any statistical test to analyze the outcome sickness absence [71]. Another study did not apply a statistical test after 6 year follow-up [51–53]. Therefore, in these studies, the effectiveness of the interventions was unknown. Within studies, different treatment effects were present for sickness absence when analyzing different subgroups, such as gender [46–49, 57–59], outcome definition RTW defined as any work or regular work [51–53], outcome measurement level (sickness absence at the moment of follow-up measurement or number of sick leave days) [72], timing of follow-up measurement [23, 31], and in- or exclusion of an outlying score [66, 67]. In addition, the active intervention in one study, was the control intervention in another study [23, 56]. Due to the heterogeneity and the limited number of studies, we were not able to generate different levels of evidence.

Study characteristics

The characteristics of the studies included are presented in Table 2. Source populations were recruited from primary health care, rehabilitation practices, social security offices, workplaces and news paper advertisements. Follow-up duration ranged from 1 month to 6 years, either tested after pretest, initial absence or injury, or after the end of treatment. Of the 17 studies, 4 were prognostic cohort studies [29, 30, 63–65] and 13 studies were RCTs [23, 31, 33, 34, 43, 46–49, 51–54, 56–59, 66–69, 71, 72].

Table I Methodological quality assessment of prognostic cohort studies and RCTs

Study	A	B	C	D	E	F	G	H	I	J	K	Sum	Quality ^c
Criteria^a													
Cohort assessment													
Sandstrom 1986a, 1986b	+	-	-	+	+	+	-	+	+			6	
Keijsers 1990	+	-	-	+	-	+	+	+	+			6	
Lindstrom/Ohlund 1992a/1992b/ 1994a/1994b/1995/1996a/1996b	+	+	-	+	+	+	-	+	+			7	
Hansen 1993	-	+	-	+	+	-	+	+	+			6	
Friedman 1995	+	-	-	+	-	+	+	+	+			6	
Loisel 1997/1998/2002	-	+	-	+	+	+	+	+	+			7	
Friedrich 1998	+	+	+	+	-	-	+	+	+			7	
Torstensen 1998	+	-	+	+	+	+	-	-	+			6	
Lonn 1999/2001	-	+	+	+	+	+	+	+	+			8	
Soukup 1999/2001	-	+	+	+	+	+	+	+	+			8	
Durand 2001	+	+	-	+	-	+	-	+	+			7	
Aure 2003	+	+	+	+	+	+	-	+	+			8	
Niemisto 2003	+	+	-	+	-	+	-	+	+			6	
Storheim 2003	+	+	+	+	-	+	+	-	+			7	
Van den Hout 2003	+	+	-	+	+	-	+	+	+			7	
Schultz 2004	-	+	-	+	+	-	+	+	+			6	
Staal 2004	-	+	-	+	-	+	+	+	+			6	
Criteria^b													
RCT assessment													
Keijsers 1990	-	-	-	-	-	-	-	-	+	+	+	3	Low
Lindstrom/Ohlund 1992a/1992b/ 1994a/1994b/1995/1996a/1996b	-	-	+	-	-	-	-	+	+	+	-	4	Low
Hansen 1993	-	-	+	-	-	+	-	-	-	+	-	3	Low
Loisel 1997/1998/2002	+	-	-	-	-	+	-	-	+	+	-	4	Low
Friedrich 1998	-	-	+	-	-	+	-	-	-	+	-	3	Low
Torstensen 1998	-	+	+	-	-	+	-	-	-	+	+	5	Low
Lonn 1999/2001	-	-	+	-	-	-	-	+	+	+	+	5	Low
Soukup 1999/2001	-	+	+	-	-	+	-	+	+	+	+	7	High
Aure 2003	+	+	+	-	-	+	+	-	+	+	+	8	High
Niemisto 2003	+	+	+	-	-	+	+	+	+	+	+	9	High
Storheim 2003	+	+	+	-	-	+	-	-	+	+	+	7	High
Van den Hout 2003	+	+	+	-	-	+	+	-	-	+	-	6	High
Staal 2004	+	+	+	-	-	+	+	+	+	+	+	9	High

^aSee Appendix 3.

^bSee Appendix 4.

^cAn RCT was considered to be of high quality if the methodological quality score for RCTs was more than 50% of the quality score (RCT score of 6 or more, with a maximum score of 9).

Predictors of outcome

Socio-demographics

Absence threshold. No socio-demographic factors were studied as predictor for sickness absence at the moment of follow-up measurement. As predictor for number of sick leave days, no evidence was found for age [63, 64] and gender [63, 64].

Table 2 Study characteristics

Reference	Design	Analysis	N	Male (%)	Sick listed at baseline (%) ^a	Source population selected from:	FU duration ^b	Outcome
Sandstrom, 1986a, 1986b	Prognostic Cohort	Correlation, fishers permutation test, logistic regression	52	67	100	Departments of Orthopedic surgery, Goteborg, Sweden	Pretest, 1 yr, 2, 4 years after pretest	1. RTW, any work or started vocational rehabilitation; 2. Days on sick leave dichotomised (<25 days and 25 days or more); 3. Days on sick leave dichotomised (< 6 months and 6 months or more) Days on sick leave during follow-up period
Keijsers, 1990	RCT	ANOVA, MANOVA	77	51	18 (43 no job)	Maastricht Back School (primary health care), the Netherlands	Baseline, 2, 6 months after assignment to treatment or control groups	Days on sick leave during follow-up period
Lindstrom, 1992a, 1992b, 1994a, 1994b, 1995, 1996a, 1996b	RCT	t-test, correlation, log likelihood ratio	103	69	100	Blue-collar workers employed at Volvo company in Goteborg, Sweden, sick listed for 6 wks because of LBP	Pretest, 1yr, 2yr; 1 yr before	1. RTW, previous job, at least half time; 2. Days on sick leave until RTW; 3. Days on sick leave dichotomised (3–6 months sick leave and > 6 months sick leave
Hansen, 1993	RCT	Wilcoxon Pratt, Kruskal-Wallis	180	68	0	Advertisements in internal company newspaper of Scandinavian Airline System	Pretest, 1, 6, 12 months after treatment termination	Days on sick leave during follow-up period
Friedman, 1995	Prognostic Cohort	Wilcoxon, logistic regression	135	76	98	Compensated LBP patients referred by their general practitioner or specialist to the rehabilitation centres	Baseline, program exit (about 4–8 weeks)	RTW, any work full or part-time
Loisel, 1997, 1998, 2002	RCT	Kaplan Meier, Hazard Ratio	104	60	100	Workplaces in the Sherbrook area (Quebec, Canada), employees on sick leave for > 4wks < 3 months (Workers Compensation Board)	Pretest, 12, 24, 52 weeks after the back incident (Loisel 1998); 12 months after initial absence, 6 yr (mean 6.4 yr, range 5.1–7.5)	1. Days on sick leave until RTW (RTW = previous job) 2. Days on sick leave until RTW (RTW = any work, full or part time); 3. (Change in) work status (not working, lighter work, regular work)
Friedrich, 1998	RCT	χ^2 , t-test and U-test	93	49	45–50	Outpatient department of orthopaedic physical therapy, Orthopaedic Hospital Speising, Vienna, Austria	Pretest, 8 th treatment session, 4, 12 months (after pretest)	RTW, previous occupational level
Torstensen, 1998	RCT	-	208	50	100	Social Security Offices in Oslo, sick listed 8–52 weeks for LBP	Pretest, posttest (3 mths), 15 mths (= 12 months after posttest)	1. RTW, not defined; 2. Days on sick leave until RTW (RTW not defined)

Author, Year	Study Design	Measures	n	n	Intervention	Follow-up	
Lonn, 1999/2001	RCT	Repeated measures ANOVA	81	46	0	Advertising in local media and referrals from other health professionals, finished treatment and sick leave Pretest, 5, 12, 36 months (after program initiation)	Days on sick leave during follow-up period
Soukup, 1999/2001	RCT	Repeated measures ANOVA	69	49	0	Advertising in local media and referrals from doctors, physical therapists, chiropractors in general clinical practice, finished treatment and sick leave Pretest, 5, 12, 36 months (after pretest)	Days on sick leave during follow-up period
Durand, 2001	Prognostic cohort	χ^2 , Fisher exact, multivariate logistic regression	127	59	100	Hospital based work rehabilitation facility in Sherbrooke, and a university hospital back pain facility in Quebec, workers absent from work, compensated 6 months after treatment	RTW, previous job
Aure, 2003	RCT	Risk Ratio	49	53	100	Social Security Office Pretest, posttest, 4 wks, 6 months, 12 months (after intervention)	% sick listed, partly or fully
Niemisto, 2003	RCT	Repeated measures ANOVA	204	46	0–9	Voluntary recruited employed CLBP patients, disability score > 16 % (Oswestry) Pretest, 5 months, 12 months (after randomisation)	Days on sick leave during follow-up period
Storheim, 2003	RCT	Kruskal Wallis, Mann-whitney, ANCOVA	93	48	100	Local national insurance offices and from general practitioners, Oslo, Norway, patients sick listed for LBP Pretest, 18 wks after inclusion	Days on sick leave during follow-up period
Van den Hout, 2003	RCT	χ^2 , logistic regression	84	76	86	Employees recently absent from work, referred to the study by general practitioners, occupational physicians, rehabilitation physicians Pretest, 6, 12 months after the intervention	1. RTW, any work full or part-time; 2. Days on sick leave during follow-up period
Schultz, 2004	Prognostic cohort	t-test + logistic regression	253	72	100	Workers compensation board of British Columbia 3 months (RTW); 18 months after injury (total days lost)	1. RTW, not defined; 2. Days on sick leave from start injury
Staal, 2004	RCT	Hazard Ratio	134	94	100	KLM Royal Dutch Airlines, absent from work because of LBP Pretest, 3 and 6 months (after randomisation)	1. Days on sick leave until RTW; 2. Days on sick leave until RTW dichotomised (< 50 days or 50 days or more); (RTW = full time for at least 4 weeks)

^aDefined as partly or on full sick leave, only defined as sick listed, or sick listed from regular tasks.

^bDiscrepancies with Table 1 are caused by timing of follow-up measurement: respectively after baseline or after finishing treatment.

Table 3 Predictors of outcome other than interventions

Reference	Socio-demographic	Life style	Medical (history)	Pain	Observed disability	Self-reported disability	Health beliefs	Physical work demands	Psychological work demands	Emotions	Expectations
Sandsstrom, 1986a, 1986b	Age; Gender; Marital status; Income; Life events	Smoking; Alcohol; Overweight; Criminality	Duration of LBP; Duration of sick leave; Analgesics; Previous healthcare utilisation; General medical history; Postural abnormalities	Previous pain in cervical and thoracic region; Pain frequency	Range of motion; Hamstring flexibility	Disability in ADL (own attitude)	-	Self-reported work demands strength; Self-reported work postures	Fatigue end of working day	-	Own expectation of recovery; Opinion rehabilitation team member; Opinion relatives
Lindstrom/ Ohlund, 1992a, 1992b, 1994a, 1994b, 1995, 1996a, 1996b	Economy; Work history; Housing; Social network; Family related problems; Life events		Analgesics; Previous healthcare utilisation; Postural abnormalities	Pain intensity; LBP complaints; Musculoskeletal complaints	Range of motion; MMH; Abdominal muscle endurance; Jump height; Arm strength; Back muscle strength; Fitness	Disability in ADL (Waddell Index, Quebec)	Physical health index	Self reported work demands strength; Vibrations; Draughts	Precision demands (self-reported and observed); Industrial work demands	Pain behavior; Illness behavior	

Friedman, 1995	Age; Gender; Work history	Duration of sick leave; Previous back surgery; Radiating pain	Pain intensity (VAS)	Range of motion; Hamstring flexibility; MMH; MVO2; Grip strength	Leisure score; Activity score; Sleep score; Self-care score	Disability in ADL (Spinoscope)	Job satisfaction	Depression (Zung); Illness behaviour; Distress
Loisel, 1997, 1998, 2002								
Durand, 2001	Age; Gender; Marital status; Educational level	Time between accident and follow-up; Previous back surgery; Radiating pain						
Van den Hout, 2003		Duration of sick leave						
Schultz, 2004		Duration of sick leave	Pain intensity (von Korff); MPQ; % bodily pain (pain drawing); SF36 bodily pain	SF36 health transition; SF36 general health; SF36 mental health	Disability in ADL (RMDQ)	Disability in ADL (RMDQ/SIP); SF36 physical functioning; SF 36 social functioning; SF36 physical component; Karasek skill discretion; Pain disability	Karasek Co-worker support; Karasek psychological demand	Depression (CES-D); State Anxiety
								Own expectation of recovery; Employer response

Legend: VAS, Visual Analogue Scale; MPQ, McGill Pain Questionnaire; MMH, Manual Materials Handling; ADL, activities of daily living; RMDQ, Roland Morris Disability Questionnaire; SIP, Sickness Impact Profile; CES-D, Center for Epidemiologic Studies Depression Scale

Table 4 Interventions as predictor for sickness absence

Study	Intervention	Operational definitions	Frequency and duration	Outcome
Absence threshold Hansen, 1993	Intensive dynamic back exercise program; Standardized physical therapy; Placebo-control	Trunk lifting, leg lifting and pull down to the neck; Obligatory and individual program. Obligatory: soft tissue treatment, manual traction, flexibility exercises, ergonomic counselling, coordination, progressive exercise program. Individual: massage, hot/cold, stretch, flexibility and stabilizing exercises, articulation; Semi-hot packs, intermittent traction Exercises (strength, endurance, coordination and stretching) and biomechanical ergonomic education. Specific attention to body awareness and the way to apply ergonomic knowledge in specific real life situations; No practical training sessions. Free to receive other treatment or exercises	All treatments 2 times a week 60 min in the course of 4 weeks	No significant difference between interventions. Significant difference within the standardized physical therapy group in number of sick leave days the year before and after treatment
Lonn, 1999, 2001	Active back school; Control	Exercises (strength, endurance, coordination and stretching) and biomechanical ergonomic education. Specific attention to body awareness and the way to apply ergonomic knowledge in specific real life situations; No practical training sessions. Free to receive other treatment or exercises	20 × 60 min in the course of 13 weeks	Active back school beneficial over control group after 1 and 3 year follow-up in number of sick leave days during follow-up
Soukup, 1999, 2001	Mensendieck; Control	Didactic and practical training (focus on applying basic ergonomic principles, strength training, stretching. Usage of Medical training therapy equipment; No practical training sessions. Free to receive other treatment or exercises	20 × 60 min in the course of 13 weeks	Mensendieck was beneficial over the control group in the 1 year follow-up (with outlier). No differences were found in the 3 year follow-up
Niemisto, 2003	Manipulative treatment; Consultation only	Manipulation using a muscle energy technique and stabilizing exercises, pain free; Information (oral and book) about LBP, how to handle with an acute phase, lifting, ADL. Free to receive other treatment or exercises	4 × 60 min in the course of 4 weeks; 2 × 60 min, pre-test and after 5 months	No differences between manipulative treatment and consultation only group in days on sick leave during follow-up period

<p>RTW threshold Lindstrom/Ohlund, 1992a, 1992b, 1994a, 1994b, 1995, 1996a, 1996b,</p>	<p>Graded activity; Traditional care</p>	<p>Including measurements of functional capacity, work place visit, back school education, individual submaximal gradually increased exercise program (operant-conditioning behavioural program) not ergonomic or other changes in work situation; Could include sick listing with rest, analgesics, available physical therapy</p>	<p>3 times a week until RTW. Duration of one session unknown</p>	<p>Graded activity was beneficial over traditional care for total group and males, both the 1 and 2year follow-up. No differences between groups in number of sick leave days for females in the 1 and 2 year follow-up</p>
<p>Loisel, 1997, 1998, 2002</p>	<p>Clinical intervention; Occupational intervention; Full intervention; Usual care</p>	<p>Visit back care school, sometimes also fitness development and work hardening with cognitive behavioural approach; Visit occupational physician and participatory ergonomics evaluation, worksite evaluation; Clinical and occupational intervention; Provided by workers physician. Free to describe any test, treatment or referral to a specialist for care</p>	<p>Unknown. Course: occupational intervention after 6 weeks of absence from work, clinical intervention after 8 weeks of absence from work</p>	<p>Occupational intervention beneficial over no occupational intervention for number of absence days. Comparing 4 intervention groups: Full intervention beneficial over usual care. Non significant difference between full versus clinical intervention, full versus occupational and occupational versus clinical intervention; Same analyses RTW any work instead of regular work: no statistically benefit in any group or combination of groups Not statistically tested</p>
<p>Tortstensen, 1998</p>	<p>Medical exercise Therapy; Conventional physiotherapy; Self-exercise</p>	<p>Graded exercise, mobilizing and stabilizing exercises. Individually designed, usage of specially designed exercise equipment; Heat or cold, massage, stretching, electrotherapy, traction, a few exercises on the treatment table; Walking</p>	<p>All treatments 36 × 60 min, 3 times a week over a course of 12 weeks</p>	<p></p>

Table 4 Continued

Study	Intervention	Operational definitions	Frequency and duration	Outcome
Durand, 2001	Therapeutic RTW program; Functional restoration therapy; Community services; Refused rehabilitation group	Work rehabilitation program is proposed to workplace management, partial tasks in agreement with occupational therapist, and progressively augmented in time and strength. Includes also functional restoration therapy, not operationalized; Not operationalized; Exclusion of any rehabilitation intervention; Referred to functional restoration and RTW intervention, but were denied by the workers compensation board	In the course of 4–8 weeks until full RTW. Not specified	Therapeutic RTW program beneficial over functional restoration and refused rehabilitation group; No differences between therapeutic RTW and community services, however, community services did not differ from Functional restoration therapy
Aure, 2003	Exercise therapy; Manual therapy	Strengthening, stretching, mobilizing, coordination and stabilizing exercises. No group training, massage and methods; Spinal manipulation, specific mobilization, stretching techniques. Pain free	All treatments 16 × 45 min, 2 times a week over a period of 8 weeks	Exercise therapy more RTW than manual therapy on the post-test, 6 month and 12 month follow up after intervention. No differences were found on the 4 week follow up
Storheim, 2003	Intensive group training; Cognitive intervention; Usual care	Norwegian aerobic fitness model. Increase overall fitness and functional capacity (cardiovascular, strength, flexibility, body awareness and relaxation.). No pain focus, safe to move; Explanation of pain mechanisms, functional examination with feedback, instruction in activation of deep stabilizing muscles and advice on how to use it actively in ADL. Instruction squat technique, how to cope with new attacks, reassure and emphasize on safe to move and use the back without restriction; Treated by general practitioner and no restrictions of treatments or referrals	2 to 3 times a week sessions of 60 min, over 15 weeks; 2 consultations were offered each lasting between 30 and 60 min. Afterwards, opportunity to contact physical therapist for advice or consultations	No differences between groups in number of sick leave days from inclusion to 18 weeks follow-up

Van den Hout, 2003	Graded activity and problem solving; Graded activity and back education	Graded activity: Operant behavioural treatment, including registration of baseline levels, treatment contract, positive reinforcement for activity increments, workplace visit, back education and lifting instructions; Problem solving: Skills training and application of skills in daily life. Group education: back and back pain related issues discussion. No skills were taught and each theme no more than one protocolized session	All groups 19 times half day sessions in a course of 8 weeks (2 to 3 times a week); Graded activity: 15 session of 60 min training; 3 additional sessions dedicated to back education and lifting instructions. 30 min with occupational therapist each week; Problem solving and group education: 10 sessions of 90 min 2 times a week 60 min until RTW, with a maximum of 3 months	No significant differences between groups on the 6 and 12 month follow-up in RTW and in number of sick leave days in the 6 month follow-up; Graded activity and problem solving less number of sick leave days than graded activity and back education in the 12 month follow-up
Staal, 2004	Graded activity; Usual care	Exercises and education messages, time contingent, regardless of level of pain; LBP guidance according to LBP Practitioners. Guidance and advice from occupational physician	2 times a week 60 min until RTW, with a maximum of 3 months	No difference in number of sick leave days in group with < 50 sick leave days. Graded activity beneficial over usual care in group of patients with 50 sick leave days or more
Threshold unknown Keijsers, 1990	Multidisciplinary back school; Waiting list group	Back school education and skills program in a group setting (10–12 patients) primarily aimed at pain management; Patients on a waiting list for multidisciplinary back school	7 times 150 min, and a refresher session after 6 months	No differences between back school and waiting list group in number of sick leave days on the 6 month follow-up
Friedrich, 1998	Motivation (exercise program and motivation program); Control (exercise program)	Exercise program: individual, submaximal gradually increased exercise program, directed towards improvement spinal mobility, muscle length, force, endurance and coordination; Motivation program: extensive counselling and information strategies. Reinforcements techniques as reward and punishment, treatment contract and keeping exercise diary were used	10 × 25 min, on average 2,3 times a week; Attached to treatments and at home (diary)	Within motivation group significant improvements in working ability over time. (pretest–4 months), not significant between pre-test and 12 months. No significant differences within control group and between intervention and control group by either 4 or 12 month follow-up

RTW threshold. As predictor for sickness absence at the moment of follow-up measurement, limited evidence was found for work history. Patients who had a job available, were working or were in training had less sickness absence at the moment of follow-up measurement [30]. No evidence was found for age [29, 30, 63, 64], gender [29, 30, 63, 64], marital status [29, 63, 64], educational level [63, 64], income, [63, 64] life events [63, 64], family related problems [63, 64] and registration in social welfare office [63, 64]. As predictor for total number of sick leave days, no evidence was found for economy [46–49, 57–59], work history [46–49, 57–59], family related problems [46–49, 57–59], life events [46–49, 57–59], housing [46–49, 57–59] and social network [46–49, 57–59].

Life style

Absence threshold. No life style factors were studied as predictor for sickness absence at the moment of follow-up measurement and as predictor for total number of sick leave days.

RTW threshold. As predictor for sickness absence at the moment of follow-up measurement, no evidence was found for smoking [63, 64], alcohol [63, 64], overweight [63, 64] and criminality [63, 64]. No life style factors were studied as predictor for total number of sick leave days.

Medical (history)

Absence threshold. No medical (history) factors were studied as predictor for sickness absence at the moment of follow-up measurement. As predictor for total number of sick leave days, no evidence was found for previous duration of sick leave [63, 64].

RTW threshold. As predictor for sickness absence at the moment of follow-up measurement, limited evidence was found for consumption of analgesics, in that patients with a low consumption of analgesics had less sickness absence at the moment of follow-up measurement [63, 64]. Conflicting evidence was found for duration of sick leave [30, 63–65] and radiating pain [29, 30]. No evidence was found for previous back surgery [29, 30], previous healthcare utilization [63, 64], general medical history [63, 64], postural abnormalities [63, 64] and time between accident and follow up [29]. As predictor for total number of sick leave days, conflicting evidence was found for previous duration of sick leave [65, 72]. No evidence was found for analgesic drug consumption, response to previous treatment and radiological findings (postural abnormalities) [46–49, 57–59].

Pain

Absence threshold. No pain factors were studied as predictor for sickness absence at the moment of follow-up measurement and as predictor for total number of sick leave days.

RTW threshold. As predictor for sickness absence at the moment of follow-up measurement, limited evidence was found for previous pain in cervical and thoracic region, in that patients with more previous pain had more sickness absence at the moment of follow-up measurement [63, 64]. No evidence was found for pain intensity [30, 65], pain frequency [63, 64], McGill Pain Questionnaire (MPQ) [65], pain drawing bodily pain [65] and Short Form 36 (SF36) bodily pain [65]. As predictor for total number of sick leave days, limited evidence was found for bodily pain [65] and musculoskeletal complaints [46–49, 57–59], in that patients with more pain or complaints had more sick leave days. Conflicting evidence was found for pain intensity [46–49, 57–59, 65]. No evidence was found for the MPQ [65], SF36 bodily pain [65], and complaints of LBP [46–49, 57–59].

Observed disability

Absence threshold. No observed disability factors were studied as predictor for sickness absence at the moment of follow-up measurement and as predictor for total number of sick leave days.

RTW threshold. As predictor for sickness absence at the moment of follow-up measurement, conflicting evidence was found for Range of Motion (ROM) [30, 51–53, 63, 64]. No evidence was found for hamstring flexibility [30, 63, 64], lifting capacity [30], grip strength [30] and aerobic capacity [30].

As predictor for total number of sick leave days, limited evidence was found for abdominal muscle endurance, in that patients with less muscle endurance had more sick leave days [46–49, 57–59]. Conflicting evidence was found for ROM, Manual Materials Handling (MMH), jump height and arm strength [46–49, 57–59]. No evidence was found for fitness and back muscle strength [46–49, 57–59].

Self reported disability

Absence threshold. No self reported disability factors were studied as predictor for sickness absence at the moment of follow-up measurement and as predictor for total number of sick leave days.

RTW threshold. As predictor for sickness absence at the moment of follow-up measurement, limited evidence was found for leisure-, activity-, and sleep score in that patients with a better leisure-, activity-, and sleep score had less sickness absence at the moment of follow-up measurement [30]. Conflicting evidence was found for disability in activities of daily living (ADL) [51–53, 63–65]. No evidence was found for SF36 physical and social functioning [65], SF36 physical component [65], Karasek skill discretion [65], pain disability [65] and self-care score [30].

As predictor for total number of sick leave days, limited evidence was found for Karasek skill discretion in that patients with less skill discretion had more sick leave days [65]. Conflicting evidence was found for self-reported disability in ADL [46–49, 57–59, 65, 72]. No evidence was found for SF36 physical and social functioning and physical component [65] and pain disability [65].

Health beliefs

Absence threshold. No health belief factors were studied as predictor for sickness absence at the moment of follow-up measurement and as predictor for total number of sick leave days.

RTW threshold. As predictor for sickness absence at the moment of follow-up measurement, limited evidence was found for health transition score in that patients with a worse health transition score had more sickness absence at the moment of follow-up measurement [65]. No evidence was found for SF36 general and mental health [65]. As predictor for total number of sick leave days, limited evidence was found for health transition score in that a worse health transition score was associated with more sick leave days [65]. No evidence was found for SF36 general and mental health [65] and the health index [46–49, 57–59].

Physical work demands

Absence threshold. No physical work demands factors were studied as predictor for sickness absence at the moment of follow-up measurement and as predictor for total number of sick leave days.

RTW threshold. As predictor for sickness absence at the moment of follow-up measurement, no evidence was found for Karasek physical demands [65] and self-reported work demands strength and postures [63, 64]. As predictor for total number of sick leave days, no evidence was found for Karasek physical demands [65], self-reported and observed work demands strength and postures [46–49, 57–59], vibrations [46–49, 57–59] and draughts [46–49, 57–59].

Psychological work demands

Absence threshold. No psychological work demands factors were studied as predictor for sickness absence at the moment of follow-up measurement and as predictor for total number of sick leave days.

RTW threshold. As predictor for sickness absence at the moment of follow-up measurement, limited evidence was found for fatigue at the end of a working day in that patients with less fatigue had less sickness absence at the moment of follow-up measurement [65]. No evidence was found for Karasek psychological demands [65], job satisfaction [30], and co-worker support [65].

As predictor for total number of sick leave days, limited evidence was found for varied work in that less varied work was associated with more sick leave days [46–49, 57–59]. No evidence was found for Karasek psychological demands [65], co-worker support [65] and industrial work demands subjectively believed to cause LBP [46–49, 57–59].

Emotions

Absence threshold. No emotional factors were studied as predictor for sickness absence at the moment of follow-up measurement and as predictor for total number of sick leave days.

RTW threshold. As predictor for sick leave at the moment of follow-up measurement, conflicting evidence was found for depression [30, 65] and no evidence was found for anxiety [65], illness behavior and distress [30]. As predictor for number of sick leave days, conflicting evidence was found for behavioral signs [46–49, 57–59], and no evidence was found for depression [65] and state anxiety [65].

Expectations

Absence threshold. No factors on expectations were studied as predictor for sickness absence at the moment of follow-up measurement. As predictor for total number of sick leave days, limited evidence was found for opinion of relatives about illness and condition in that more negative expectations were associated with more sick leave days [63, 64]. No evidence was found for own expectations [63, 64].

RTW threshold. As predictor for sickness absence at the moment of follow-up measurement, consistent evidence was found for own expectations of recovery in that patients with higher expectations of recovery had less sickness absence at the moment of follow-up measurement [63–65]. Limited evidence was found for recommendation of team rehabilitation member in that patients with a positive recommendation of the rehabilitation team member had less sickness absence at the moment of follow-up measurement [63, 64]. No evidence was found for opinion of relatives [63, 64]. As predictor for total number of sick leave days, limited evidence was found for own expectations and employer response in that lower own expectations and a lower employer response were associated with more sick leave days [65].

Discussion

In summary, for the absence threshold, no predictors were found for factors predicting sickness absence at the moment of follow-up measurement, and no consistent evidence was found for predictors for total number of sick leave days, because predictors were only studied once. Aimed at the RTW threshold, consistent evidence was found for own expectation of recovery as predictor for sickness absence at the moment of follow-up measurement. Patients with higher expectations of recovery had less sickness absence at the moment of follow-up measurement. No consistent evidence was found for predictors for total number of sick leave days. Due to the heterogeneity and the limited number of studies, we were not able to generate different levels of evidence for intervention strategies as predictor for sickness absence aimed at the absence threshold or RTW threshold. It can be concluded that no core set of predictors exists for sickness absence in general. The characteristics of the study (the absence threshold and RTW threshold, study population, timing of follow-up measurement, predictors and outcome definition) determined the influence of several predictors on sickness absence in patients with CLBP.

Only the factors age, gender and opinion of relatives were studied for both the absence threshold and RTW threshold. For both thresholds, no evidence was found for age and gender. For the absence threshold, a worse opinion of relatives was associated with more sick leave days, but for the RTW threshold, no evidence was found for opinion of relatives as predictor for sickness absence at the moment of follow-up measurement. For the RTW threshold, several predictors were studied for sickness absence at the moment of follow-up measurement and number of sick leave days. Evidence was found that more pain intensity, more bodily pain and less skill discretion was associated with more sick leave days, but no evidence was found for these factors as being predictive for sickness absence at the moment of follow-up measurement. In addition, much conflicting evidence was present. This can be explained in that predictors were measured using different instruments or definitions, which might not be comparable. For example, previous duration of sick leave was either dichotomized in 4–6 weeks and 6–12 months, or used continuously as number of sick leave days. Pain intensity was measured with the VAS and the Von Korff scale, self-reported disability in ADL measured with the Oswestry, RMDQ, SIP, Quebec, Waddell disability index and attitude to own ADL capacity, and depression was measured with the Zung and the CES-D. In addition, radiating pain, ROM and MMH are compound scores. The separate items may not be similarly predictive, i.e. more lifting capacity might be associated with a shorter duration of sickness absence, but pushing and pulling may not be associated with sickness absence. Then, it would not be legitimate to analyze the compound scores as predictors for sickness absence. However, because not all studies presented separate items, we decided to analyze the compound scores also. In addition, within studies, differences were found in classifications of subgroups. With respect to definition of RTW, being male was facilitating for fulltime RTW, but not for any RTW. With respect to definition of timing of follow-up measurement, more behavioral signs were associated with more sick leave days in the one-year follow-up, but not in the two-year follow-up. With respect to gender, less jump height and less arm strength were associated with more sick leave days in males, but not in the total group (males and females).

Previous reviews also showed that no specific set of predictors for sickness absence in patients with CLBP can be found. In addition, the reviews also mentioned that studies were very heterogeneous with respect to study population, predictors and outcomes [10–13]. Similar to our review, no consistent evidence was found for the outcome participation restriction (which includes RTW) in the most recently performed review [11], which means that a limited number of studies existed that studied the same predictive factors. In addition, previous studies showed that predictors of outcome may vary with the definitions used [73, 74]. In our review, we

confirmed this. The cause of the lack of set of predictors can be explained in that studies are too heterogeneous to summarize, even for qualitative analyses.

To gain insight in the predictive factors for sickness absence, we recommend the development of a core set of measurements in the evaluation of CLBP and the use of uniform definitions of outcome measurements. A first attempt to propose uniform definitions of LBP has already been given [75]. However, no core set of measurements exist yet and although recommendations of a minimum data-set to assess work status was given [76], a uniform definition of RTW is still lacking. We recommend using a definition of RTW that is based on the Dutch social security laws: RTW defined as full return to regular work with a minimum duration of 6 weeks. Regular work is defined as the previous job or new (temporary) job with similar work demands. Recurrent episodes of absence from work because of LBP should be considered as belonging to the first continuous period of absence from work. In addition, total number of days absence from work should be registered as initial days (in a group already sick listed) and as days of recurrent episodes of absence from work associated with LBP. If patients partly RTW or perform modified duties with lower work demands, they remain on the sick list, and thus every day accounts for a sick listing day.

Distinction should be made between predictors aimed at the absence threshold and at the RTW threshold. Therefore, it is recommended to study the subgroup working at the beginning of the study and the already sick listed, or when studying a mixed population, it should be described to what extent workers remain sick listed, what percentage of workers deteriorate (from working to sickness absence) or recover (from sickness absence to RTW), so that the difference between the absence and RTW threshold can be studied.

CLBP is mostly traditionally defined as low back pain over 12 weeks of duration [77]. This review however, included patients with LBP over 4 weeks duration (Appendix 2) and also patients with intermittent LBP over a longer period of time, in which the current episode may last shorter than 4 weeks. Therefore, studies may be included in this review, which were excluded in previous studies, but also different predictors of outcome may be found compared to other reviews, because of the use of a different definition for CLBP [73]. However, the above mentioned population was selected for inclusion in this review because of the intermittent character of LBP [75, 78] and because the importance of reduction of sickness absence already starts in the transition stage from acute to chronic LBP.

Bias

Selection of studies

Although a thorough literature search was performed, publication bias cannot be excluded. It is possible that only studies were published that generated positive results on the outcome, disregarding studies that generated negative outcomes or no evidence. Papers were included if the study population was defined as patients with LBP or musculoskeletal pain with a subgroup of LBP presented separately. However, because LBP was used as search term instead of musculoskeletal complaints, it is possible that studies on musculoskeletal complaints exist, that also present a subgroup of LBP in the text of the paper that were not retrieved from our literature search. However, if important outcomes were present in those studies, we assumed that this would be described in the abstract. Therefore we also searched for LBP as free text word in title or abstract.

Quality of the studies

All criteria were assessed by two independent reviewers, one content expert and a non-expert with a methodological background, as is recommended to prevent bias by prior opinions [18]. Of the 50 prognostic cohort studies, 17 studies were of high quality for prognostic studies. Of the 17 prognostic studies, 13 studies were RCTs. Despite papers selected on the basis of a high quality score for prognostic studies, 7 low quality RCTs were found. This demonstrates the different approach of quality assessment between RCTs and observational studies. The RCT quality list is widely used and recommended by the Cochrane Collaboration, developed to prevent bias in generating the evidence [18]. No widely accepted quality criteria are available for assessing the methodological quality of prognostic studies. Therefore, the quality list applied in this study was based on criteria defined by the Cochrane Collaboration for observational studies [15], Borghouts [16] and Scholten Peeters [17] (Appendix 4). Selection bias may have occurred, because the choice of the criteria for quality assessment may have influenced the final inclusion in the review. When analyzing the individual criteria for cohort studies, it was observed that all studies clearly described potential prognostic factors and the performed analyses adequately. However, only 5 of the 17 studies [23, 31, 33, 33, 54, 66, 67, 69] described both their in- and exclusion criteria adequately, as well as characteristics of their study population (criteria B & C, Appendix 3). The lack of adequate description of study population may be a potential source of clinical heterogeneity. When analyzing individual RCT criteria, it was observed that none of the studies blinded patients and care providers. This was expected because most interventions were exercise treatments. Mostly, patients and caregivers cannot be blinded from exercises. It is noticeable that of the included studies, the studies performed most recently all have high quality. This might be due to the fact that recently performed RCTs follow the Cochrane Collaboration Guidelines. This may indicate bias, because the guidelines should be seen as a state of the art for reviews and not as some kind of gold standard [18]. Perhaps previous RCTs did follow guidelines in their study design, but failed to report their study adequately. In addition, different papers of the same study cohort were analyzed as was one study for quality assessment. The more papers published, the higher the chance of a high quality score, because only one of the papers should have described the criteria adequately. This may lead to a higher quality than if the papers were assessed individually. All quality criteria were assessed for the total study. During the scoring of the studies, it was observed that some criteria were assessed as negative for the total study, but when we assessed the criteria according to our outcome measure, it should be scored positive. For example, number of dropouts was not acceptable for the total population, but of all patients' sick leave data were obtained at the end of the study. Therefore, for sick leave data, the number of dropouts was acceptable. Some studies that were excluded for review because of low quality, might have been included when focusing the quality assessment on our outcome criteria instead of on the total study.

Assessment of evidence

The intention of this review was to overcome heterogeneity by selecting papers according to predetermined criteria for prognostic quality to achieve clinically homogeneous studies that might enable statistical pooling. Pooling of data is only relevant and meaningful if studies are comparable on study characteristics as study population, predictive factors/ intervention strategies and outcome measures. Otherwise, pooling of studies will result in systematically biased estimates [79, 80]. Few studies fulfilled the criteria, therefore we decided to use a cut-off score for inclusion, to select only high quality studies and as a consequence, instead of pooling, a qualitative analysis of the included studies was performed. However, due to the heterogeneity

of the studies, we demonstrated that qualitative summary might also lead to biased estimates. It could be argued that if a predictive factor was associated with sickness absence but only studied once, whether this should be defined as limited evidence. In a previous review [81], availability of 1 study was defined as no evidence instead of limited evidence. It was argued that consistency of evidence could not be evaluated on the basis of one study. However, we decided to define it as limited evidence, because we already selected the studies on high quality for prognostic cohort studies, therefore, a certain level of evidence is assured.

Conclusion

It can be concluded that no core set of predictors exists for sickness absence in general and that the characteristics of the study (the decision to report sick or to return to work, study population, timing of follow-up measurement, predictors and outcome definition) determine the influence of several predictors on sickness absence in patients with CLBP. This also means that subgroups of patients may exist within the CLBP population, which should be treated differently to achieve desirable outcomes such as return to work. Until now, too few studies are available to overcome several potential sources of heterogeneity and to investigate and compare predictors of outcome for different subgroups. Therefore, the evidence presented in this review should be used with caution, due to the unknown influence of other potential sources of heterogeneity. Further research and use of a core set of measurements and uniform definitions are needed to predict the decision to report sick or to return to work in patients with CLBP.

Appendix 1: Literature search

Each database was searched for: “Low Back Pain” entered as MesH term and free text word (Psychinfo only LBP in title or abstract), combined with MesH terms and free text words. In each database, the same free text words were used. Mesh Terms differed per database. In Table A1 the MesH terms are presented. Some MesH terms were only used in combination with the free text words “work” or “working” or “occupation*” or “job*” or “employment” or “unemployment,” presented in the table as combination terms.

Free text words used for the searches:

Return* to work	Off work	Work disability
Job resumption	Absent from work	Work ability
Job loss	Back to work	Work incapacity
Work loss	Work status	Work capacity
Work resumption	Employment status	Occupational disability
Work absence	Occupational status	Occupational ability
Absenteeism	Vocational status	Occupational incapacity
Sick* leave*	Job status	Occupational capacity

Appendix 2: In- and exclusion criteria

A paper was included when all inclusion criteria were met and none of the exclusion criteria were applicable; a paper was included for further judgment when the exclusion criteria did not apply, but (some of) the inclusion criteria were not clearly specified. A paper was excluded when one of the exclusion criteria were met.

Table A1 Mesh terms used in the different databases

Medline	AMED	Embase	Cinahl	Psychinfo	Cochrane
Sick leave Absenteeism Work (explode) Work capacity evaluation	Sick leave Absenteeism Work capacity evaluation	Absenteeism Work Work capacity Job performance	Sick leave Absenteeism Work Work capacity evaluation Job re-entry	Employee absenteeism Job performance Occupational status	Sick leave Absenteeism Work Work capacity evaluation
Rehabilitation vocational (explode) Occupational diseases/rehabilitation Employment (explode)	Rehabilitation vocational (explode) Employment (explode)	Work resumption Work disability Vocational rehabilitation Occupational diseases/rehabilitation Employment Employability	Rehabilitation vocational Occupational diseases/rehabilitation Employment status Reemployment	Vocational rehabilitation Employment status (explode)	Rehabilitation vocational (explode) Employment (explode)
Combination terms ^a Treatment outcome Disability evaluation Predictive value of tests	Treatment outcome Disability evaluation Predictive value of tests Clinical assessment scales Rehabilitation Disability	Treatment outcome Predicting and forecasting Disability Physical disability Social disability	Treatment outcomes Disability evaluation Predictive value of tests Clinical assessment tools	Treatment outcome Disability evaluation Statistical validity	Treatment outcome Disability evaluation Predictive value of tests

^aThese combination Mesh terms were only used in combination with the free text words “work” or “working” or “occupation*” or “job*” or “employment” or “unemployment”

Inclusion criteria

I1 Study population	
1	Non-specific low back pain (LBP) or musculoskeletal pain in which a subgroup of patients with LBP is presented separately. Non-specific = pain (with or without radiation) without specific origin. Low back = the lumbar spine
2	Sub acute, chronic or recurrent LBP Sub acute – chronic = complaints and/or sick leave associated with back pain \geq 4 weeks Recurrent LBP = defined as ‘recurrences or episodes of back pain in the previous year’
I2 Design	
3	Observational study, (prospective cohort study, follow-up study, longitudinal study) or Randomized Control Trial (with therapy as prognostic factor)
4	Prognostic factors should be identified (including age, gender)
I3 Outcome	
5	In a cohort of patients on sick leave at baseline: <ul style="list-style-type: none"> – RTW during study yes/no – Duration of sick leave (in days) In a cohort of patients working at baseline: <ul style="list-style-type: none"> – Sick leave during study yes/no – Duration of sick leave (in days) – Number of sick leave registrations

In a mixed population (on sick leave and working) both outcomes can be applicable. These outcomes should be described separately.

Exclusion criteria

E1 Study population	
a	Only a group of patients with musculoskeletal complaints other than LBP
b	LBP with a distinct causal diagnose (Bechterew, Rheumatoid Arthritis, spondylolisthesis, fracture, infection, inflammatory process, neoplasm) or a mixed population without making the distinction between non-specific and specific LBP. Degeneration of the spine without nerve compression (spondylosis) is not a specific cause for LBP
c	Patients with cardiovascular or pulmonal disease, hypertension, drug addiction, psychopathology or pregnancy, or a mixed population in which no distinction is made between these groups and the non-specific LBP group
d	Patients in which >10% has a post surgery status in the past 2 years, or a mixed population in which no distinction is made between post surgery patients and non surgery LBP patients
e	Duration of LBP or sick leave \leq 4 weeks, or a mixed population in which acute, sub acute and chronic patients were not described separately
f	Duration of complaints = ‘New workers compensation claim’
E2 Design	
g	Cross-sectional studies without follow-up period, retrospective studies, (systematic) reviews, updates of (systematic) reviews, abstracts of congress papers, commentary on other papers/letters to the editor

h	Study population < 40 LBP patients (case studies, case reports)
i	No prognostic factors analyzed (including age, gender)
E3 Outcome	
j	Outcome other than described above

Appendix 3: Methodological criteria list for cohort studies

A. Was the duration of complaints \geq 6 weeks?		Yes/No/?
B. Was the description of in- and exclusion criteria adequate? (1 or 2, and 3)	1 Inclusion 2 Exclusion 3 Pregnancy	Yes/No/?
C. Was the study population/potential prognostic factors described adequately? (all 5)	Age Gender Duration complaints Sick leave Surgery	Yes/No/?
D. Are other potential prognostic factors described adequately? (at least 1)	1 Physical 2 Psychosocial 3 Work related 4 Pre-existing 5 Financial 6 Socio-demographic 7 Treatment	Yes/No/?
E. Was the follow-up duration \geq 12 months?		Yes/No/?
F. Was the number of dropouts described and acceptable?		Yes/No/?
G. Are demographic/ clinical data described adequately of the dropouts? (at least 1)	Description Statistical test	Yes/No/?
H. Was the outcome defined adequately?		Yes/No/?
I. Are the performed analyses adequately (described)? (1 or 2 or 3 and 4)	1 e.g. T-test 2 e.g. Risk ratio 3 e.g. Regression analysis 4 Correction for multiple comparisons	Yes/No/?

Appendix 4: Methodological criteria list for assessing RCTs (18)

A	Was the method of randomization adequate?	Yes/No/?
B	Was the treatment allocation concealed?	Yes/No/?
C	Were the groups similar at baseline regarding the most important prognostic indicators?	Yes/No/?
D	Was the patient blinded to the intervention?	Yes/No/?
E	Was the care provider blinded to the intervention?	Yes/No/?
F	Was the outcome assessor blinded to the intervention?	Yes/No/?
G	Were co-interventions avoided or similar?	Yes/No/?

H	Was the compliance acceptable in all groups?	Yes/No/?
I	Was the dropout rate described and acceptable?	Yes/No/?
J	Was the timing of the outcome assessment in all groups similar?	Yes/No/?
K	Did the analysis include an intention-to-treat analysis?	Yes/No/?

Acknowledgements The authors want to thank M.F. Reneman (PhD) for his comments on a previous version of this manuscript. This study was supported by grants of 'Zorgonderzoek Nederland', number 96-06-006, the Foundation 'Beatrixoord Noord Nederland' and the Foundation 'De Drie Lichten', the Netherlands.

References

1. Waddell G, Main CJ. A new clinical model of low back pain and disability. The back pain revolution. London: Churchill Livingstone; 1998. p. 223–240.
2. Deyo RA, Rainville J, Kent DL. What can the history and physical examination tell us about low back pain? *JAMA* 1992;268:760–765.
3. Andersson GBJ. Epidemiological features of chronic low-back pain. *Lancet* 1999;354:581–585.
4. Waddell G. The physical basis of back pain. In: Waddell G, editor. The back pain revolution. London: Churchill Livingstone; 1998. p. 135–154.
5. Kendall NAS, Linton SJ, Main CJ. Guide to assessing psychosocial yellow flags in acute low back pain: risk factors for long-term disability and work loss. Wellington, New Zealand; 1997.
6. Nachemson AL. Newest knowledge of low back pain. A critical look. *Clin Orthop* 1992;279:8–20.
7. Maetzel A, Li L. The economic burden of low back pain: a review of studies published between 1996 and 2001. *Best Pract Res Clin Rheumatol* 2002;16:23–30.
8. Allegro JT, Veerman TJ. Sickness absence. In: Drenth JD, Thierry H, de Wolff CJ, editors. Handbook of work and organizational psychology. East Sussex: Psychology Press; 1998. p. 121–144.
9. Philadelphia Panel evidence-based clinical practice guidelines on selected rehabilitation interventions for low back pain. *Phys Ther* 2001;81:1641–1674.
10. Elders LA, van der Beek AJ, Burdorf A. Return to work after sickness absence due to back disorders: a systematic review on intervention strategies. *Int Arch Occup Environ Health* 2000;73:339–348.
11. van der Hulst M, Vollenbroek-Hutten MMR, IJzerman MJ. A systematic review of sociodemographic, physical and psychological predictors of multidisciplinary rehabilitation – or, back school treatment outcome in patients with chronic low back pain. *Spine* 2005;30:813–825.
12. Crook J, Milner R, Schultz IZ, Stringer B. Determinants of occupational disability following a low back injury: a critical review of the literature. *J Occup Rehabil* 2002;12:277–295.
13. Truchon M, Fillion L. Biopsychosocial determinants of chronic disability and low-back pain: a review. *J Occup Rehabil* 2000;10:117–142.
14. Waddell G, Burton AK. Occupational health guidelines for the management of low back pain at work: evidence review. *Occup Med* 2001;51:124–135.
15. Etiology and prognosis: critical appraisal form for observational studies. Cochrane Collaboration website 2002; Available from: URL: <http://www.cochrane.dk/nrsmg>
16. Borghouts JA, Koes BW, Bouter LM. The clinical course and prognostic factors of non-specific neck pain: a systematic review. *Pain* 1998;77:1–13.
17. Scholten-Peeters GGM, Verhagen AP, Bekkering GE, van der Windt DAWM, Barnsley L, Oostendorp RAB, Hendriks EJM. Prognostic factors of whiplash-associated disorders: a systematic review of prospective cohort studies. *Pain* 2003;104:303–322.
18. van Tulder M, Furlan A, Bombardier C, Bouter L, editorial board of the Cochrane Collaboration Back Review Group. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine* 2003;28:1290–1299.
19. Assendelft WJJ, Scholten RJPM, Van Eijk JTHM, Bouter LM. De praktijk van systematische reviews. III. Methodologische beoordeling van onderzoeken. *Ned Tijdschr Geneesk* 1999;143:714–719.
20. Verhagen AP, de Vet HCW, de Bie RA, Boers M, van den Brandt PA. The art of quality assessment of RCTs included in systematic reviews. *J Clin Epidemiol* 2001;54:651–654.
21. Coté P, Cassidy JD, Carroll L. A systematic review of the prognosis of acute whiplash and a new conceptual framework to synthesize the literature. *Spine* 2001;26:E445–E458.
22. Bigos SJ, Bowyer O, Braen GI. Acute low back problems in adults. Clinical Practice Guideline No. 14. AHCPR Publication No. 95-0642. Rockville, MD: U.S Department of Health and Human Services; 1994.

23. Aure OF, Nilsen JH, Vasseljen O. Manual therapy and exercise therapy in patients with chronic low back pain: a randomized, controlled trial with 1-year follow-up. *Spine* 2003;28:525–531.
24. Bachman S, Oesch PR, Kool JP, Persili S, Knusel O. Treatment of patients with chronic low back pain in a functional restoration program: Work Related function parameters, pain parameters and the working status after 12 months. *Physikalische Medizin Rehabilitationsmedizin Kurortmedizin* 2003;13:263–270.
25. Bendix T, Bendix A, Labriola M, Haestrup C, Ebbelohj N. Functional restoration versus outpatient physical training in chronic low back pain: a randomized comparative study. *Spine* 2000;25:2494–2500.
26. Bentsen H, Lindgarde F, Manthorpe R. The effect of dynamic strength back exercise and/or a home training program in 57-year-old women with chronic low back pain. Results of a prospective randomized study with a 3-year follow-up period. *Spine* 1997;22:1494–1500.
27. Beurskens AJ, de-Vet HC, Koke AJ, Regtop W, van der Heijden GJ, Lindeman E, Knipschild PG. Efficacy of traction for nonspecific low back pain. 12-week and 6-month results of a randomized clinical trial. *Spine* 1997;22:2756–2762.
28. Casso G, Cachin C, Van Melle G, Gerster JC. Return-to-work status 1 year after muscle reconditioning in chronic low back pain patients. *Joint Bone Spine* 2004;71:136–139.
29. Durand MJ, Loisel P. Therapeutic return to work: Rehabilitation in the workplace. *Work: J Prev, Assess Rehabil* 2001;17:57–64.
30. Friedman PJ, Leadley MJ, Stickney J, Austin KL. Prediction of return to work following rehabilitation for chronic low back injury. *N Z J Occup Ther* 1903;46:20–24.
31. Friedrich M, Gittler G, Halberstadt Y, Cermak T, Heiller I. Combined exercise and motivation program: effect on the compliance and level of disability of patients with chronic low back pain: a randomized controlled trial. *Arch Phys Med Rehabil* 1998;79:475–487.
32. Gibson T, Grahame R, Harkness J, Woo P, Blagrove P, Hills R. Controlled comparison of short-wave diathermy treatment with osteopathic treatment in non-specific low back pain. *Lancet* 1985;1:1258–1261.
33. Glomsrod B, Lonn JH, Soukup MG, Bo K, Larsen S. “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.
34. Hansen FR, Bendix T, Skov P, Jensen CV, Kristensen JH, Krohn L, Schioeler H. Intensive, dynamic back-muscle exercises, conventional physiotherapy, or placebo-control treatment of low-back pain. A randomized, observer-blind trial. *Spine* 1993;18:98–108.
35. Hazard RG, Fenwick JW, Kalisch SM, Redmond J, Reeves V, Reid S, Frymoyer JW. Functional restoration with behavioral support. A one-year prospective study of patients with chronic low-back pain. *Spine* 1989;14:157–161.
36. Hildebrandt J, Pffingsten M, Saur P, Jansen J. Prediction of success from a multidisciplinary treatment program for chronic low back pain. *Spine* 1997;22:990–1001.
37. Hurley DA, McDonough SM, Dempster M, Moore AP, Baxter GD. A randomized clinical trial of manipulative therapy and interferential therapy for acute low back pain. *Spine* 2004;29:2207–2216.
38. Hurri H. The Swedish back school in chronic low back pain. Part I. Benefits. *Scand J Rehabil Med* 1989;21:33–40.
39. Jousset N, Fanello S, Bontoux L, Dubus V, Billabert C, Vielle B, Roquelaure Y, Penneau-Fontbonne D, Richard I. Effects of functional restoration versus 3 hours per week physical therapy: a randomized controlled study. *Spine* 2004;29:487–493.
40. Karjalainen K, Malmivaara A, Mutanen P, Pohjolainen T, Roine R, Hurri H. Outcome determinants of subacute low back pain. *Spine* 2003;28:2634–2640.
41. Karjalainen K, Malmivaara A, Pohjolainen T, Hurri H, Mutanen P, Rissanen P, Pahkajarvi H, Levon H, Karpoff H, Roine R. Mini-intervention for subacute low back pain: a randomized controlled trial. *Spine* 2003;28:533–540.
42. Karjalainen K, Malmivaara A, Mutanen P, Roine R, Hurri H, Pohjolainen T. Mini-intervention for subacute low back pain: two-year follow-up and modifiers of effectiveness. *Spine* 2004;29:1069–1076.
43. Keijsers JF, Steenbakkens MW, Meertens RM, Bouter LM. The efficacy of the back school: a randomized trial. *Arthritis Care Res* 1990;3:204–209.
44. Kendrick D, Fielding K, Bentley E, Miller P, Kerslake R, Pringle M. The role of radiography in primary care patients with low back pain of at least 6 weeks duration: a randomised (unblinded) controlled trial. *BMJ* 2001;322:400–405.
45. Licciardone JC, Stoll ST, Fulda KG, Russo DP, Siu J, Winn W, Swift J. Osteopathic manipulative treatment for chronic low back pain: a randomized controlled trial. *Spine* 2003;28:1355–1362.
46. Lindstrom I, Ohlund C, Eek C, Wallin L, Peterson LE, Fordyce WE, Nachemson AL. The effect of graded activity on patients with subacute low back pain: a randomized prospective clinical study with an operant-conditioning behavioral approach. *Phys Ther* 1992;72:279–290.

47. Lindstrom I, Ohlund C, Eek C, Wallin L, Peterson LE, Nachemson A. Mobility, strength, and fitness after a graded activity program for patients with subacute low back pain. A randomized prospective clinical study with a behavioral therapy approach. *Spine* 1992;17:641–652.
48. Lindstrom I, Ohlund C, Nachemson A. Validity of patient reporting and predictive value of industrial physical work demands. *Spine* 1994;19:888–893.
49. Lindstrom I, Ohlund C, Nachemson A. Physical performance, pain, pain behavior and subjective disability in patients with subacute low back pain. *Scand J Rehabil Med* 1995;27:153–160.
50. 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;22:1610–1616.
51. Loisel P, Abenham L, Durand P, Esdaile JM, Suissa S, Gosselin L, Simard R, Turcotte J, Lemaire J. A population-based, randomized clinical trial on back pain management. *Spine* 1997;22:2911–2918.
52. Loisel P, Poitras S, Lemaire J, Durand P, Southiere A, Abenham L. Is work status of low back pain patients best described by an automated device or by a questionnaire? *Spine* 1998;23:1588–1594.
53. Loisel P, Lemaire J, Poitras S, Durand MJ, Champagne F, Stock S, Diallo B, Tremblay C. Cost-benefit and cost-effectiveness analysis of a disability prevention model for back pain management: a six year follow up study. *Occup Environ Med* 2002;59:807–815.
54. Lonn JH, Glomsrod B, Soukup MG, Bo K, Larsen S. Active back school: prophylactic management for low back pain. A randomized, controlled, 1-year follow-up study. *Spine* 1999;24:865–871.
55. Moffett JK, Torgerson D, Bell-Syer S, Jackson D, Llewlyn-Phillips H, Farrin A, Barber J. Randomised controlled trial of exercise for low back pain: clinical outcomes, costs, and preferences. *BMJ* 1999;319:279–283.
56. Niemisto L, Lahtinen-Suopanki T, Rissanen P, Lindgren KA, Sarna S, Hurri H. A randomized trial of combined manipulation, stabilizing exercises, and physician consultation compared to physician consultation alone for chronic low back pain. *Spine* 2003;28:2185–2191.
57. Ohlund C, Lindstrom I, Areskou B, Eek C, Peterson LE, Nachemson A. Pain behavior in industrial subacute low back pain. Part I. Reliability: Concurrent and predictive validity of pain behavior assessments. *Pain* 1994;58:201–209.
58. Ohlund C, Eek C, Palmbald S, Areskou B, Nachemson A. Quantified pain drawing in subacute low back pain. Validation in a nonselected outpatient industrial sample. *Spine* 1996;21:1021–1030.
59. Ohlund C, Lindstrom I, Eek C, Areskou B, Nachemson A. The causality field (extrinsic and intrinsic factors) in industrial subacute low back pain patients. *Scand J Med Sci Sports* 1996;6:98–111.
60. Petersen T, Kryger P, Ekdahl C, Olsen S, Jacobsen S. The effect of McKenzie therapy as compared with that of intensive strengthening training for the treatment of patients with subacute or chronic low back pain: a randomized controlled trial. *Spine* 2002;27:1702–1709.
61. Pflingsten M, Hildebrandt J. Treatment of chronic low back pain through intensive activation—an assessment of 10 years. *Anesthesiol Intensivmed Notfallmed Schmerzther* 2001;36:9–589.
62. Pflingsten M, Hildebrandt J, Saur P, Franz C, Seeger D. Multidisciplinary treatment program on chronic low back pain, part 4. Prognosis of treatment outcome and final conclusions. *Schmerz* 1997;11:1–41.
63. Sandstrom J, Esbjornsson E. Return to work after rehabilitation. The significance of the patient's own prediction. *Scand J Rehabil Med* 1986;18:29–33.
64. Sandstrom J. Clinical and social factors in rehabilitation of patients with chronic low back pain. *Scand J Rehabil Med* 1986;18:35–43.
65. Schultz IZ, Crook J, Meloche GR, Berkowitz J, Milner R, Zuberbier OA, Meloche W. Psychosocial factors predictive of occupational low back disability: towards development of a return-to-work model. *Pain* 2004;107:77–85.
66. Soukup MG, Glomsrod B, Lonn JH, Bo K, Larsen S. The effect of a Mensendieck exercise program as secondary prophylaxis for recurrent low back pain. A randomized, controlled trial with 12-month follow-up. *Spine* 1999;24:1585–1591.
67. Soukup MG, Lonn J, Glomsrod B, Bo K, Larsen S. Exercises and education as secondary prevention for recurrent low back pain. *Physiother Res Int* 2001;6:27–39.
68. Staal JB, Hlobil H, Twisk JW, Smid T, Koke AJ, Van Mechelen W. Graded activity for low back pain in occupational health care: a randomized, controlled trial. *Ann Intern Med* 2004;140:77–84.
69. Storheim K, Brox JI, Holm I, Koller AK, Bo K. Intensive group training versus cognitive intervention in sub-acute low back pain: short-term results of a single-blind randomized controlled trial. *J Rehabil Med* 2003;35:132–140.
70. Storro S, Moen J, Svebak S. Effects on sick-leave of a multidisciplinary rehabilitation programme for chronic low back, neck or shoulder pain: comparison with usual treatment. *J Rehabil Med* 2004;36:12–16.
71. Torstensen TA, Ljunggren AE, Meen HD, Odland E, Mowinckel P, Geijerstam S. Efficiency and costs of medical exercise therapy, conventional physiotherapy, and self-exercise in patients with chronic low back pain. A pragmatic, randomized, single-blinded, controlled trial with 1-year follow-up. *Spine* 1998;23:2616–2624.

72. Van den Hout JHC, Vlaeyen JWS, Heuts PHTG, Zijlema JHL, Wijnen JAG. Secondary prevention of work-related disability in nonspecific low back pain: does problem-solving therapy help? A randomized clinical trial. *Clin J Pain* 2003;19:87–96.
73. Ozguler A, Leclerc A, Landre M, Pietrie-Taleb F, Niedhammer I. Individual and occupational determinants of low back pain according to various definitions of low back pain. *J Epidemiol Community Health* 2000;54:215–220.
74. Wasiak R, Verma S, Pransky G, Webster B. Risk factors for recurrent episodes of care and work disability: case of low back pain. *J Occup Environ Med* 2004;46:68–76.
75. De Vet HCW, Heymans MW, Dunn KM, Pope DP, van der Beek AJ, Macfarlane GJ, Bouter LM, Croft PR. Episodes of low back pain. A proposal for uniform definitions to be used in research. *Spine* 2002;27:2409–2416.
76. Amick BC, Lerner D, Rogers WH, Rooney T, Katz JN. A review of health-related work outcome measures and their uses, and recommended measures. *Spine* 2000;25:3152–3160.
77. Frymoyer JW. Back pain and sciatica. *N Engl J Med* 1988;318:291–300.
78. Von Korff M, Saunders KJD. The course of back pain in primary care. *Spine* 1996;21:2833–2837.
79. Scholten RJPM, Kostense PJ, Assendelft WJJ, Bouter LM. de praktijk van systematische reviews. IV. Het combineren van de resultaten van afzonderlijke onderzoeken. *Ned Tijdschr Geneesk* 1999;143:786–791.
80. Scholten RJPM, Assendelft WJJ, Kostense PJ, Bouter LM. De praktijk van systematische reviews. V. Heterogeniteit tussen onderzoeken en subgroepanalyses. *Ned Tijdschr Geneesk* 1999;143:843–848.
81. Hoogendoorn WE, van Poppel MNM, Bongers PM, Koes BW, Bouter LM. Physical load during work and leisure time as risk factors for back pain [review]. *Scand J Work Environ Health* 1999;25:387–403.