Eur Spine J (2007) 16:1755–1775
DOI 10.1007/s00586-007-0433-8

REVIEW

Value of predictive instruments to determine persisting restriction of function in patients with subacute non-specific low back pain. Systematic review

Roger Hilfiker · Lucas M. Bachmann · Carolin A.-M. Heitz · Tobias Lorenz · Harri Joronen · Andreas Klipstein

Received: 1 June 2007 / Accepted: 18 June 2007 / Published online: 15 August 2007

Abstract Low back pain (LBP) can restrict function with all the personal, interpersonal, and social consequences, such as a loss of independence and the inability to fulfil diverse roles in social life. Therefore, the prevention of the consequences of LBP would reduce costs, individual burdens and social burdens. Being able to fulfil the requirements of daily living is a cornerstone of quality of life. Early identification of patients who are likely to develop chronic pain with persistent restricted function is important, as effective prevention needs informed allocation of health care and social work. The aim of this study was to report and discuss the predictive value of instruments used to identify patients at risk of chronic LBP. Medline, Embase, CINAHL, Central, PEDro, Psyndex, PsychInfo/PsycLit, and Sociofile were systematically searched up to July 2004. Reference lists of systematic reviews on risk factors, and reference lists of the studies included were also searched. The selected studies evaluated predictive values of tools or predictive models applied 2–12 weeks after an initial medical consultation for a first or a new episode of non-specific LBP with restriction in function. Instruments had to predict function-related outcomes. Because of the heterogeneity of the instruments used we did not pool the data. Sixteen publications on function-related outcomes were included. The predictive instruments in these studies showed only moderate ability to predict or explain function-related outcome (maximal 51% of the variability). There was great variability in the predictors included and not all known risk factors were included in the models. The reviewed tools showed a limited ability to predict function-related outcome in patients with risk of chronic low back pain. Future instruments should be based on models with a comprehensive set of known risk factors. These models should be constructed and validated by international, coordinated research teams.

Keywords Back pain · Prognosis · Function

RH conceived the study, drafted the protocol, reviewed the literature, selected the literature, extracted data, analysed data and drafted the manuscript. TL conceived the study, reviewed the literature and assisted draft of the manuscript. LMB conceived the study, drafted the protocol and added substantial revisions on manuscript. CH and HJ reviewed and selected the studies. AKL conceived the study and added substantial revisions on manuscript. All authors read and approved the final manuscript.

R. Hilfiker
Domain Health and Social Work, Physiotherapy,
University of Applied Sciences Western Switzerland,
Valais, Leukerbad, Switzerland

L. M. Bachmann
Horten Centre, University of Zurich, Zurich, Switzerland

C. A.-M. Heitz · H. Joronen · A. Klipstein
Department of Rheumatology
and Institute of Physical Medicine,
University Hospital Zurich, Zurich, Switzerland

T. Lorenz
Rehazentrum Leukerbad, Leukerbad, Switzerland

R. Hilfiker · C. A.-M. Heitz
Physiotherapy Sciences Studies, Department of Rheumatology
and Institute of Physical Medicine, University Hospital Zurich,
Switzerland and Faculty of Health Sciences,
Maastricht University, Maastricht, The Netherlands

R. Hilfiker (✉)
Swiss Paraplegic Research (SPF), 6207 Nottwil, Switzerland
e-mail: roger.hilfiker@hevs.ch; roger.hilfiker@paranet.ch

Background

Non-specific back pain has a life time prevalence of about 80% [44]. More than half of the patients with non-specific low back pain (LBP) resume full functioning within one month after a new onset of a LBP episode; 75–90% of the patients resume full functioning within 3 months [30]. Performing activities of daily living is one of the main components of quality of life, and is, besides return to work, the most important patient-centred outcome. Return to work as a sole outcome is not sufficient as, (a) one will exclude many patients who did not work before the LBP episode (housewives, elderly people), (b) return to work depends strongly on external factors (e.g. economic situation in the given region or profession), and (c) return to work and function are strongly correlated, but do not represent the same underlying construct. Therefore, it is important to evaluate prediction of function in a separate analysis. Persistent pain can restrict function, with consequences not only for the patient him/herself in different life situations, e.g. interpersonal interactions, but also with consequences for those close to the patient in work or private life.

Recent guidelines recommend the assessment of risk factors for severe disease and prolonged disability if a patient with LBP has not returned to full activity at 4 and 6 weeks after the onset of disabling LBP [21, 22, 34, 40]. Predicting or explaining functional recovery or disability would help to concentrate precious health care and social work on patients in need, especially by an informed assignment of the different interventions. If one knows a patient's modifiable risk factors for persistent limitations, an informed assignment to different interventions is possible. For example, if social risk factors are predominant over medical factors, emphasis may be put on social work. To disclose a broader set of risk factors than those traditionally accounted for from the medical perspective would enhance the effectiveness of health care providers. In addition, explanation would be provided for ineffective medical interventions in the case that social risk factors are predominant.

Furthermore, these predictive tools could be used to "negotiate" and explain the rationale for management strategies to the patients. This might enhance the patient's understanding of the problem and thus, e.g., enhance compliance during the rehabilitation process.

The possibility of accurate prediction of risk in clinical trials will allow the implementation of strategies in these trials to balance out known prognostic factors or to control for confounders in the analyses [31].

Several authors systematically reviewed the risk factors for function-related outcomes [7, 8, 15, 26, 29, 32, 35, 37, 45], but to our knowledge, there is no systematic overview of instruments predicting persisting restriction of function for patients with non-specific LBP.

We systematically reviewed the literature to find predictive models and tools for the transition from subacute to chronic non-specific LBP with persistent restriction in function. To report and discuss the predictive value of the instruments found, we evaluated the methodological quality, discriminative properties, and ability to predict or to explain the function-related outcome of these studies of patients with subacute non-specific LBP.

Methods

The search for evidence to answer the question "value of predictive tools to determine long-term restriction in function" was combined with the search for the similar study question "value of predictive tools to determine long-term non-return to work". In this publication we report only the values for the predictive instrument for persisting restriction of function.

We searched studies reporting predictive values of questionnaires, assessments, clinical examination, etc.), or models (combining different individual risk factors or assessments to a "decision rule", "clinical rule" or "predictive-tool") for the prediction of chronic non-specific LBP with persisting restriction in function. For simplicity of reading, we will use the term "instruments" to summarize all these assessments and clinical rules, etc. in this text.

Inclusion criteria were: prospective cohort study, patients with subacute non-specific LBP and instruments had to be applied between 2 and 12 weeks after the initial medical visit for a first or a new episode of LBP [45]. We excluded retrospective studies, studies that applied the predictive instruments in a general population, studies that applied the instruments at a too early time-point (less than 10 days) or too late (more than 3 months) after the medical visit because of an onset of a new LBP episode, studies that included pregnant patients, patients with neck pain or patients with specific pathologies such as inflammatory diseases, cancer or studies that did not have at least a three-month follow up.

An epidemiologist and an information specialist defined the search strategy (available from the first author) for the different electronic databases. The search had no language, date, or publication status restriction. These systematic searches were conducted in July 2004: Medline in-Process (Ovid version, 1966–2004), Embase, (1974–2004) PsychINFO/PsychLIT (1987–2004), CINAHL (1982–2004), Central (2nd Quarter 2004, PEDro (from inception to 2004), Psyndex (1977–2004), Sociofile (1974–2004). In addition, we checked reference lists of the publications

included of relevant systematic reviews, relevant articles on the topic, guidelines and expert reports. Furthermore, we searched in the Related Articles section of the studies included and reviews in PubMed (also after July 2004). Studies for the research questions were selected in two stages: initially, two reviewers (RH, TL) independently assessed 55% of the retrieved abstracts; two other reviewers (CH, HJ) assessed 55% of the identified abstracts. Agreement for the overlapping 10% was analysed and judged as good. In a second step, two reviewers (CH, RH) read the full text of the pre-selected studies and used checklists to decide on definite inclusion. Articles in other languages were assessed by physiotherapists with sufficient knowledge of the given language. Disagreements were resolved by discussion with the second author (LMB).

We checked the methodological quality of each study with a criteria checklist based on recommendations [11, 12, 18, 19] (see legend Table 2 for items). As effects of individual quality components can be masked by simple summary scores, we additionally looked for serious methodological flaws that would strongly bias the predictive values and which were not covered by the check-list [17]. The following additional issues were evaluated: number of events, whether interaction was considered, data reducing methods (for example principle component analyses) and variable selection process. Special issues for multivariable models were discussed [5]. We extracted all available relevant data at baseline, and the univariable and multivariable associations and predictive values. We extracted as much relevant data as needed and as possible to answer the question of predictive value. Given the heterogeneity in the tools and models, a meta-analysis would not have provided meaningful interpretable information.

Results

Selection process

Figure 1 shows the flowchart of the inclusion process of the combined search for work related outcome (not reported here) and function-related outcome. We identified 4,968 references in the databases and finally included 15 publications on function-related outcomes [4, 9, 10, 13, 14, 20, 23, 25, 27, 36, 39, 42, 43, 46, 47].

Table 1 shows the studies included. Publication years ranged from 1993 to 2005. Five studies used logistic regression [4, 20, 27, 39, 42], six studies multiple linear regressions [10, 13, 14, 23, 36, 47], one study used latent transition regression analyses [43], two studies used recursive partitioning [9, 10], one study used analysis of covariance [25] and one study used univariable analyses only [46].

Quality

See Table 2 for rating of the items from the checklist. All studies were classified as having moderate quality. The checklist does not reflect the general quality of the study, merely the quality concerning predictive issues. Percentages of patients available at follow up ranged from 56 to 100%. Five reports had follow-up rates below 80% [4, 13, 14, 36, 43] (see Table 1).

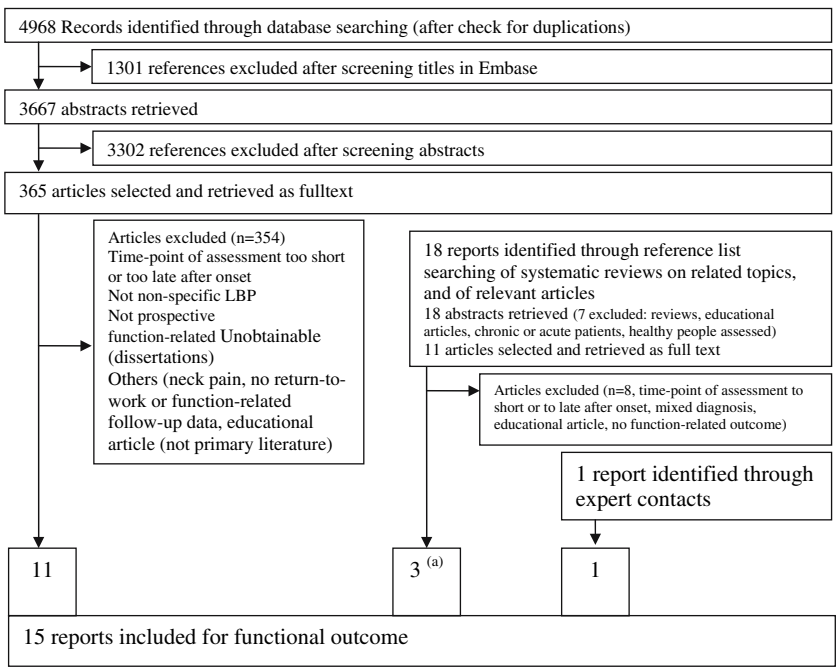

Fig. 1 In this article we report only on studies on function-related outcomes. (*a*) Note that two articles were published after the end of the systematic search of the databases [36, 43], and one article was excluded in the abstract screening at first due to the wording "..in patients with chronic low back pain..." in the title [14]

Table 1 Included studies dealing with the transition from subacute to chronic low back pain

	First author, country, year				
	Carey [4], USA, 2000	Dionne [10], Canada, 1997 Training sample and validation sample	Dionne [9], Canada, 2005	Epping [13], Jordan, USA, 1998 Same population as in Williams [47] and Wahlgren [46]	Hansson [14], Denmark, Germany, Israel, Netherlands, Sweden, United States, 2000
N[a], rate at baseline, rate at follow up	1639, ?, 76% (at 22 months)	1213, 72%, 84% Training sample (TS): 569, 83% Validation sample (VS): 644, 85%	860, 68%, 85% Validation study of the final model built in Dionne 1997 [10]	138, ?, 55.7%	Denmark: 565,?, 87% at 1 year, 78% (at 2 years) Germany: 410,?, 72%, 63% Israel: 327,?, 88%,81% Netherlands: 427; 92%,87% Sweden: 539; 84%,73% US: 484,?, 85%, 79%
Study participants	208 physicians, chiropractors and mid-level practitioners at a group-model Health Maintenance Organization. No data available for the missing 24%. Data for the 76% at 22 months follow up: Mean age 42.5 (functionally better group at 12 weeks), 44.6 (others), 44% male; 24% sciatica at baseline, 83% employed at baseline, 30% on workers compensation at baseline, 17% first time LBP.	Health Maintenance Organization setting TS: Mean age 46.7 (SD 14.3), 69.6 % over 12 years education Not all sick-listed (only 3.5% received financial compensation at that time) VS: Mean age 46.5 (SD 14.8), 69.3% over 12 years education, Low socioeconomic status was underrepresented.	Primary care setting Quebec Mean age 38.7, 58.5% male. Only 1.7% were non-white. 44.3 % "only high school'', 72.1% married or living as married, 38% current smokers.	Systematic inclusion of patients in a closed healthcare system (Naval Medical Centre), 100% men, mean age 31.9 (SD 7.0)	Patients on job with sickness and/or disability benefits. Sick listed 100%,age 18 to 59 years,
Independent variables: Instruments used	Pain (0–10), gender, race, income, first episode of LBP, duration of LBP before initial visit, employment status, workers compensation, Roland Morris, sciatica, care received, medication used, satisfaction with care	Job factors (employment status at one month, full or part-time student, physical demands of the job (reaching, handling), responsibility for housekeeping, years employed at current job, D.O.T[b].). Psychosocial factors (coping catastrophizing score[c]), self-reported health status, stress from family pressures in the past six months, (SCL-90 Depr./ Somat. [d]). Injury factors (financial compensation, injury diagnosis). Medical factors (migraine, baseline disability (Roland-Morris score), self-reported type of episode, years since onset of back pain, Grouped chronic pain score[e], nr. of associated diagnoses, nr. of pain complaints, present intensity of pain, worst intensity of pain in the past 6 months (VAS), subject ever kept from full-time work for back pain, self-reported injury diagnosis, chest pain in the past 6 month, number of surgery for back pain). Demographic factors (Gender, age, height, BMI, ethnic, marital status, education, income).	SCL-Depression, SCL-Somatization[d]	Descriptor Differential Score (DDS[g]), Sickness Impact Profile (SIP[h]), Beck Depression Inventory[i]	Hannover ADL[j], Von Korff pain intensity scale[k], SF.36: Vitality, mental health, social functioning and general health[l]. Karasek Theorell's demand-support-control scale[m], age, gender, native language, treatment received before sick report, psychological demand, decision latitude (control), physical job demands, surgery during the first year, manipulation/traction, heat and cold (treatment), massage, TENS, physical therapy, back-school program, acupuncture,

Table 1 continued

	First author, country, year				
	Carey [4], USA, 2000	Dionne [10], Canada, 1997 Training sample and validation sample	Dionne [9], Canada, 2005	Epping [13], Jordan, USA, 1998 Same population as in Williams [47] and Wahlgren [46]	Hansson [14], Denmark, Germany, Israel, Netherlands, Sweden, United States, 2000
Dependent variables: Instrument used	Modified Roland Disability Scale at 3 months and 22 months.	Modified Roland-Morris at 2-years.	Functional status (Roland-Morris Disability Questionnaire) at 2 years	6 and 12 months pain intensity, 6 and 12 months disability, 6 and 12 month depressive symptoms,	Work status at 1 year and 2 years Disability (Hannover ADL) at 1 year and 2 years
Definition of non-specific LBP	LBP, no previous care for the current episode, no previous spine surgery, no non-skin malignancy, not pregnant at the time, and able to speak English. Patients were assessed within 2 weeks after medical visit, and after four weeks.	Eligible subjects were adult patients aged 18–75 years who consulted on of the HMOs primary care physicians for a back pain episode. Subjects enrolled at GHC[f] for less than one year and those who visited walk in care facilities and emergency rooms were excluded.	Eligible if back pain was non-specific and patient had caused at least 1 day of absence from job. Excluded were: pregnancy, serious comorbidity, only cervical pain, or specific cause.	First onset of back pain (thoracic 6 or below), present on a daily basis for 8 (+-2) weeks.	LBP, excluded were: back surgery during the preceding 12 months, back problems caused by fractures, inflammation, infection or tumours
Risk assessment time-point	Only the predictors assessed at 4 weeks were used for this review.	4–6 weeks after the index visit	3 weeks after their index medical consultation	6–10 weeks after onset of first time LBP	After 90 days on sick-leave
Outcome statistic variables selection process	Logistic regression	Multiple regression for continuous outcome, stepwise, minimum significance level set at 0.01 Recursive partitioning technique for the dichotomized (50%) modified Roland-Morris	Recursive partitioning with the model developed in Dionne 1997	Hierarchical Multiple Regression Analysis	Multiple linear regression (Disability)

	Karjalainen [20], Finland, 2003	Klenerman [23], UK, 1995	Leroux [25], Canada, 2004 RAMS-prognosis study	Loisel [27], Canada, 2002	Truchon [36], Canada, 2005
N[a], rate at baseline, rate at follow up	134, 81%, 97% to 98%	162, ?, 100%	1007, 68%, 84%	104, ?, 87%	439, 49%, 73%
Study participants	From 36 primary health care centres, 350 general practitioners. Mean age 44 (SD 8.8, range 25 to 60), 58% women, 22% blue-collar workers, satisfaction with work (0 to 10): 7.3 (SD 2.3) 34% with high school diploma, 69% with radicular symptoms below the knee	Patients from general practitioners (gender for this subgroup not available)	Workers seeking medical consultation in primary care setting. Baseline characteristics only for patients for which follow up was available: 483 men (mean age 39.4; SD 10.3); 366 women (mean age 38.4; SD 10.9)	Recruited from all workplaces with more than 175 workers and located within of 30 km from the back clinic. Randomized study, 53,9% men, mean age 40.2 (SD 9.1). Mean days of absence from work 40.6 (SD 14.1)	Workers on sick leave by the Quebec Workers' Compensation Board Mean age 39 years (range 18 to 60) 56% men, 63 % married or part of a couple, 60% had high school diploma, 25% higher diploma

Table 1 continued

	Karjalainen [20], Finland, 2003	Klenerman [23], UK, 1995	Leroux [25], Canada, 2004 RAMS-prognosis study	Loisel [27], Canada, 2002	Truchon [36], Canada, 2005
Independent variables: Instruments used	Gender, age, BMI, high school diploma, blue-collar worker, duration of sick leave at baseline, radicular symptoms below the knee, intensity of pain at baseline (0–10), Oswestry Disability Index, perceived risk of not recovering (0–10), expectation regarding the effectiveness of treatment (0–10), satisfaction with work (0–10), self-rated health status for age	Age, sex, referring doctor, marital status, employment status, smoking habits. Previous and present history and severity of LBP, ratings of the four fear-avoidance contextual variables: stressful life events with rating scale Holmes and Rahe, Somatic symptoms (MSPQ[n]), Previous Pain History[o], Pain Coping Strategies (Slade)[p]. Pain Drawing, Modified Zung, Disability (Oswestry) Non-organic symptoms, non-organic signs, body mass index, neurological tests, SLR[q], Prone knee bend, Hip flexion, lateral flexion, sagittal movement, area affected, clinical diagnosis by doctor.	Age, marital status, education, work schedule, duration of work in same job, stressful life events in past 12 months, baseline Roland-Morris score, self-reported type of episode, psychological demands and job decision latitude (Karasek's Job Content Questionnaire[r]), somatization score, depression score, self-reported health status, fear-avoidance beliefs about back pain, likelihood of loosing job within 2 years, psychological job demands, worst intensity of back pain in past 6 month, physical workload (level of physical effort x lifting heavy load, whole-body vibration, frequency of trunk flexion, lateral bending) Combinations of psychological demands and social supports at work (work APGAR[s]), job satisfaction	Quebec Task Force Classification (QTFC[t]).	Coping (CPCI[u]), Disability (Roland-Morris disability questionnaire), Pain (NRS-101[v]), Depressive Mood (Hospital Anxiety and Depression Scale HADS), Catastrophizing (Catastrophizing scale of Coping Strategies Questionnaire)
Dependent variables: Instrument used	High Oswestry Disability Index (dichotomized) Intensity of Pain (0 to 10), daily symptoms, bothersome Pain, Pain interfering with work or daily Life, HRQL (15D), satisfaction with medical care (0 to 10), health care cost, sick leave (0, 1 to 30, or over 30 days). At 12 months.	Level of pain, Disability, whether off work, course of LBP at 12 month For this review, only the analyses for disability are considered (because only these predict from the 2 months assessment)	Roland-Morris score at 1 year	Sickness Impact Profile (SIP[h]), Oswestry Questionnaire, McGill Pain Questionnaire (MPQ) Work status, assessed from patient questionnaire (return to the identical work performed before onset of back pain).	Pain, Disability, Depressive mood at 6 months, work status
Definition of non-specific LBP	LBP (with or without sciatica), excluding: need for operative treatment, pregnancy, history of specific back disease (cancer, fracture, spondylarthritis, or infection), somatic or psychiatric disorder preventing rehabilitation, substance abuse, consultation with a specialist in physical and rehabilitation medicine during the past year, inpatient rehabilitation for back pain during the last 3 years, 3 month of continuous sick leave during the preceding year, and impossibility of a work visit.	Benign, musculoskeletal LBP.	Non-specific back pain (including the thoracic, lumbar, and lumbosacral areas), at least 1 day of sick leave. Exclusion: specific pathologies, such as cancer, spinal infection, vertebral fractures, systematic diseases, cauda equine syndrome, referred pain, pregnancy, and major medical illness that could affect work status.	First four categories of the Quebec Task Force Classification (QTFC[t]). Pregnant, spine fracture, significant degenerative spine disease (spondylolisthesis, Grade 2 or more), non-mechanical spine disease (tumour or infection), or major comorbid condition were excluded.	On sick-leave between 3 and 12 weeks after the accident. Exclusion criteria: insufficient understanding of French, return to work, pregnancy, previous back surgery, severe spinal pathology like fracture, tumour, infection, cauda equine syndrome, and symptoms suggesting nerve compression.

Table 1 continued

	Karjalainen [20], Finland, 2003	Klenerman [23], UK, 1995	Leroux [25], Canada, 2004 RAMS-prognosis study	Loisel [27], Canada, 2002	Truchon [36], Canada, 2005
Risk assessment time-point	After 4–12 weeks of LBP making working difficult	For this review, we included only analysis from the 2 months assessment	About 3 weeks after index medical visit	4–12.5 weeks after initial medical visit	Between 3 and 12 weeks on sick leave
Outcome statistic variables selection process	Generalized estimating equations (GEE) (repeated measures data, with mean response modelled as a logistic regression model).	Principal component analyses Multiple linear regression	Analyses of Covariance with adjustment of the baseline Roland-Morris score and adjustment for confounding factors with multiple regression.	Logistic regression	Hierarchical multiple regression
	Van der Weide [39], Netherlands, 1999	Von Korff [42], USA, 1993	Von Korff [43], USA, 2005 same population as Von Korff 1993 und Dionne 1997	Wahlgren [46], USA, 1997 Same population as in Epping-Jordan 1998 and Williams 1998	Williams [47], USA, 1998 Same population as in Epping-Jordan 1998 and Wahlgren 1997
N[a], rate at baseline, rate at follow up	120, 85%, 90%	1213, 72%, 94.4%	1213, 72%, 93% (1 year), 84% (2 year)%, 68% (5 year)	138, ?%, .55%	138, %, 60%
Study participants	Patients on sick leave, health care and university workers. 33% male, mean age 39 (SD 8.7). Smokers 33%, sporting activities more than one hour per week 68%, suspicion of nerve root compression 24%, radiating pain until knee 24%, beyond knee 38%, on sick leave during last year 31%, on sick leave more than once 44%, mentally demanding work 25%, mixed mentally or physically 53%, physically demanding work 22%	Health Maintenance Organization (HMO), age 18 to 74, 47.1% male, 52.9 female, 33.2% college graduate, 61.6% high school graduate, 5.2 less than 12 years of education	Health Maintenance Organization (HMO), age 18 to 74, 47.1% male, 52.9 female, 33.2% college graduate, 61.6% high school graduate, 5.2 less than 12 years of education	Naval Medical Center, San Diego, 100% men, mean age 30.5 (SD7.3), mead education 12.8 (SD 1.8) years, 69% married 52% pain without radiation, 18% pain with proximal radiation, 13% with distal radiation, 13% pain with radiation and neurological signs, 2 patients with compression of spinal nerve root confirmed by specific imaging techniques.	as above
Independent variables: Instruments used	Demographic factors, perceptions of work conditions, low-back pain characteristics, pain intensity (VAS[w]), functional disability (Roland Disability Questionnaire), general health perception (Nottingham Health Profile), coping, health locus of control	Demographic, pain (characteristic pain intensity=average of three 0 to 10 ratings: back pain at the time of the interview, average pain, and worst pain during previous 6 months), Disability score, Disability days, Days in pain, recentness of onset (time since the first episode of back pain), Pain grade (five categories). SCL-90-R depression and vegetative symptoms scale.	Demographic, pain (characteristic pain intensity=average of three 0 to 10 ratings, back pain at the time of the interview, average pain, and worst pain during previous 6 months), Disability score, Disability days, Days in pain, recentness of onset (time since the first episode of back pain), Pain grade (five categories). SCL-90-R [d]depression and vegetative symptoms scale, number of pain sites	Descriptor Differential Scale (DDS[g]), Sickness Impact Profile (SIP[h]), Beck Depression Inventory (BDI[i])	Job satisfaction (Job Descriptive Index), modified work APGAR[s], Pain VAS, Descriptor Differential Scale[g], Sickness Impact Profile[h], Quality of Well Being Index, Hamilton Rating Scale for Depression, Beck Depression Inventory[i], Automatic Thoughts Questionnaire[w], Orthopedic impairment[x].
Dependent variables: Instrument used	Roland-Morris Disability Questionnaire at 3 and 12 months. Return to work status at 3 month, 6 month and 12 month, Sick leave at 1 year, Time to return to work.	Poor outcome (Grade III: high disability-moderately limiting back pain, Grade IV: high disability-severely limiting back pain). At 1 year	Poor outcome (Grade III: high disability-moderately limiting back pain, Grade IV: high disability-severely limiting back pain). At 1 year and at 5 years	Pain at 6 months Disability at 6 months Distress at 6 months Overall Clinical Outcome at 6 months.	Pain at 6 months Disability at 6 months Distress at 6 months Overall Clinical Outcome at 6 months.

Table 1 continued

	Van der Weide [39], Netherlands, 1999	Von Korff [42], USA, 1993	Von Korff [43], USA, 2005 same population as Von Korff 1993 und Dionne 1997	Wahlgren [46], USA, 1997 Same population as in Epping-Jordan 1998 and Williams 1998	Williams [47], USA, 1998 Same population as in Epping-Jordan 1998 and Wahlgren 1997
Definition of non-specific LBP	LBP for at least 10 days, pain located below the scapulae and above the gluteal fold. Not pregnant.	Patients at a primary care back pain visit, visits to emergency room and walk-in care facilities were not included.	Patients at a primary care back pain visit, visits to emergency room and walk-in care facilities were not included.	First onset back pain (T6 or below) present *on a daily basis* for the previous 8 (±) weeks. Excluded were prior episodes of back or other pain on a daily basis lasting one week or longer; taking medications known to affect mood (e.g. antidepressants, anxiolytics), major surgery within the preceding 12 months, and back pain secondary to neoplastic disease, osteomyelitis, of fracture.	First-onset LBP Exclude were: prior episode of back or other pain on a daily basis lasting for 2 weeks or longer, major medical illness (e. g, insulin-dependent diabetes, chronic obstructive pulmonary disease, taking medications known to affect mood (e. g, antidepressants), prior back surgery, pain secondary to neoplastic disease, osteomyelitis, or fracture).
Risk assessment time-point	After 10 days on sick-leave	3–6 weeks after index medical visit	3–6 weeks after index medical visit	6–10 weeks after onset of first time LBP	6–10 weeks after onset of first time LBP
Outcome statistic variables selection process	Logistic regression All variables with $P < 0.15$ in the univariable analysis were included in the multivariable analysis. Also included were variables with practical relevancy for the occupational physician. Interaction terms were evaluated	Logistic regression	Latent Transition Regression Analysis (LTRA) Risk score estimation	Change from baseline to six months: cross-tabulation	Principal Components Analysis, Hierarchical multiple regression

[a] Sample size. ? means not reported, *LBP* low back pain. [b] D.O.T.:Dictionary of Occupational Titles of the US Department of Labor. [c] Coping Catastrophizing Score: they used a score based on selected items from the Vanderbilt Pain Management Inventory and the Coping Strategy Questionnaire. [d] SCL-90 (Depression / Somatization): Dionne et al. (1997[10]) used selected items from the SCL-90 Depression and Somatization subscales. [e] Grouped Chronic Pain scores: pain grade based on characteristic pain intensity, disability score and number of pain disability days in the past six months). [f] GHC: Group Health Cooperative of Puget Sound, Health Maintenance Organization (HMO). [g] DDS: The Descriptor Differential Scale is a self-report measure of current pain intensity and unpleasantness. [h] SIP: The Sickness Impact Profile is a 136 self-report questionnaire that evaluates the degree to which pain interferes with usual daily activities. [i] BDI: Beck Depression Inventory, 21-Item self-administered questionnaire that assesses severity of depressive symptoms. [j] Hannover ADL: evaluates function limitations due to LBP (12 items), score from 0 (greatest possible limitations) to 100 (no limitations).[k] Von Korff pain intensity scale: score between 0 (no pain) to 10: [l] Vitality, mental health, social functioning and general health, subscales from SF-36, a health-related quality-of-life instrument. The scores for all the subscales range between 0 and 100 (higher values = better function). [m] Karasek Theorell's demand-support-control scale: A questionnaire evaluating "job strain", psychological demands, conflicts at work, and the possibility to utilizing an individuals abilities and the possibility to develop new skills. [n] MSPQ: Modified Somatic Perception Questionnaire, 13 Items for somatic symptoms. [o] Previous Pain History: scale for previous pain, 1 for no pain at all, 10 for worst imaginable pain, three categories of pain: internally produced pain (headaches, toothache, etc.), externally produced pain (fractures, dentistry, etc.) and accidental pain (cut finger etc.). [p] Pain Coping Strategies: five choices of strategies (Medication, rest, doctor, physical exercise, ignoring). [q] SLR: Straight Leg Raise. [r] Karasek's Job Content Questionnaire: 9 Items for psychological job demands, 9 Items for job decision latitude. [s] Work APGAR Scale: adaptation, partnership, growth, affection and resolve. 17 Items to evaluate the perception of support at the workplace and job enjoyment (in the Leroux study the score was dichotomized by the median). [t] Quebec Task Force Classification (QTFC): a commonly used classification based on simple clinical criteria including signs and symptoms (pain and neurological examination data), imaging test results, and response to treatment (in this study only QTFC categories 1–4 were used; 1 = Pain without radiation, 2 = Pain with proximal radiation (above knee), 3 = Pain with distal radiation (below knee), 4 = Pain with distal radiation and neurological signs. [u] CPCI: self-report measure, rates frequency of use of behavioural and cognitive coping strategies (guarding, resting, asking for assistance, relaxation, task persistence, exercise/stretching, coping self-statements, and seeking social support). Four answer categories-version used. [v] NRS-101: Numeric Rating Scale 0 to 100. [w] VAS: Visual Analogue Scale. [w]Automatic Thoughts Questionnaire: 30 statements, score is reflective of distressing cognitions and are related to sub-clinical and clinical depression. [x]Orthopaedic impairment: Waddel Impairment Index for measuring the severity of orthopaedic disease in back disorders. Score from 0 to 50 (more impairment)

[Tr]: "training sample" (to define the model), [Va]: "validation sample": They randomly divided the population in two samples: one to develop the model, the second to validate the model

Table 2 Quality of reviewed studies

Quality criteria Author	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Carey 2000	y	y	y	?	y	n	n	n	?	y	y	y	n	n	n
Dionne 1997	y	?	y	n	y	y	n	n	?	y	y	y	y	y	n
Dionne 2005	y	y	y	n	y	n	y	n	?	y	y	y	y	y	n
Epping-Jordan 1998	y	y	y	y	y	y	y	n	?	y	y	y	n	n	n
Hansson 2000	n	?	n	?	y	y	y	n	?	y	y	y	p	n	n
Karjalainen 2003	y	?	y	?	y	?	y	y	?	y	y	y	y	n	n
Klenerman 1995	y	?	y	?	y	n	n	n	?	y	y	y	y	n	n
Leroux 2004	y	?	p	n	y	y	y	n	?	y	y	y	y	n	n
Loisel 2002	y	y	y	?	n	n	n	y	?	n	y	y	y	n	n
Truchon 2005	y	?	y	n	y	n	y	y	?	y	y	y	y	y	n
Van der Weide 1999a	y	?	y	y	y	y	n	y	y	?	y	y	y	y	n
Van der Weide 1999b	y	?	y	y	y	y	n	y	y	?	y	y	y	y	n
Von Korff 2003	y	y	y	n	y	n	n	n	?	y	y	y	y	n	n
Von Korff 2005	y	y	y	n	y	n	n	n	?	y	y	y	y	n	n
Wahlgren 1997	y	y	y	y	y	y	y	n	?	y	y	y	n	n	n
Williams 1998	y	y	y	y	y	y	y	n	?	y	y	y	n	n	n

Items *1* Was the hypothesis/aim/objective of the study clearly described (prognostic)? *2* Were the patients enrolled consecutive?

3 Were the main characteristics of the included patients in the study clearly described? *4* Was the response rate at baseline at least 80% of the possibly eligible patients? *5* Were the psychosocial data collected with validated instruments? *6* Were data on physical workload collected?

7 Was a clear definition of non-specific low back pain used? *8* Was the treatment standardised?

9 Were prognostic factors that were assessed addressed by treatment? *10* Statistical adjustment for important prognostic factors?

11 Were the statistical methods adequately described? *12* Was the outcome clearly defined? *13* Were the outcome measures available for at least 80% of the included patients? *14* Was the model cross validated in a group of patients different from the group in which it was derived, preferably with different clinicians? *15* Was there a serious methodological flaw not covered by the check-list?

y yes; *n* no; *p* partially, ? means not reported/not clear

The majority of the studies did not include all relevant risk factors (e.g. specific self-efficacy or expectations of the patient regarding recovery) of the relevant domains (biomedical, psychosocial, socio-economic, as well as occupational). For example, several studies evaluated no risk factor from the psychological domain [4, 20, 27, 46]. No factors from the social and occupational domain were evaluated in several other studies [13, 27, 36, 46]. Some studies built their model from a comprehensive set of risk factors (i.e. from all relevant domains) [10, 14, 23, 25]. See Table 1 for independent variables included.

Predictive values for function-related outcomes

Table 3 shows the predictive values of the studies evaluating prediction of continuous function-related outcomes (disability questionnaires). The variance explained (R^2) in outcome, where reported, ranged from 28 to 51%; mean 42% (SD 8). The highest explained variance was observed for disability as a predictor and for psychosocial factors.

Table 4 shows the results from the studies reporting odds ratios for dichotomized functional outcome: median odds ratio (we inversed odds ratios if they were less than 1) for individual factors was 2.20 (interquartile range 1.49 to 3.68). Odds ratios above four were observed for the following predictors: lack of energy (9.9), high disability with severely limiting back pain (8.1), high disability with moderately limiting back pain (6.1), pain radiating below the knee combined with neurological signs (5.7), high score in the Oswestry Disability Index (5.2), social isolation (4.3) and avoidance coping style (4.1).

Table 5 reports probabilities of the different classifications for having impaired function at follow up. Probabilities for not having impaired function at follow up if being in the given "low risk" group ranged from 0 to 11.7% with a mean of 6% (SD 4.3). Probabilities for having impaired function at follow up if being in the "high risk" groups ranged from 26.8 to 82.1% with a mean of 51.6% (SD 18.9). Predictive values for not having impaired function are better than for predicting impaired function. For example, "not being distressed" was a better predictor of not having impaired function than "being distressed" was a predictor of having impaired function at follow up [10].

Table 3 Studies reporting explained variance from regressions. Predictive instruments and models used for function-related outcomes at different time-points. Models predict restricted function

Study	Predictive factors/instruments	R^2	*P*	ΔR^2	ΔP	β	Outcome
Dionne 1997 [10]	SCL-90-R Somatization	0.175					Modified Roland-Morris at two years
Patients with 1-month Roland-Morris score >0 in the training sample (*N* at follow up: 408)	SCL-90-R Depression	0.218					
	1-month Roland-Morris	0.262					
	Number of pain days in the past 6 months (*N*=395)	0.283					
Dionne 1997 [10]	α					–0.003203	Modified Roland-Morris at two years
total sample	SCL-90-R Somatization					0.040177	
N at follow up: 1213	SCL-90-R Depression					0.093217	
	1-month Roland-Morris					0.256079	
	Number of pain days in the past 6 months					0.000581	
Epping-Jordan 1998 [13]	2-month disability	0.45 / 0.32			0.000/0.000	0.5/0.42	6/12 months disability (SIP)
N at follow up: 78	Income (–)	0.48/0.38			0.174/0.085	–0.2/–0.7	
	Ethnicity (minority ethnicity) (– for 12 months)					0.11/–0.2	
	Age (– for 6 months)					–0.9 / 0.8	
	2-month pain intensity (– for 12 months)	0.51/0.40			0.148/0.348	0.07/–0.6	
	2 month depression					0.20/0.16	
Hansson 2000 [14], Denmark	Age		0.000				Back function (Hannover ADL) 12 months (multiple regression)
	Gender (female)		0.000				
N at follow up: 494	Physical therapy		0.010				
Hansson 2000 [14], Germany	Age		0.000				Back function (Hannover ADL) 12 months (multiple regression)
	Psychological demand		0.007				
N at follow up: 295	Heat and cold (treatment)		0.039				
Hansson 2000 [14], Israel	Age		0.000				Back function (Hannover ADL) 12 months (multiple regression)
N at follow up: 289							
Hansson 2000 [14], Netherlands	Age		0.000				Back function (Hannover ADL) 12 months (multiple regression)
	Gender (female)		0.000				
N at follow up: 392	Decision latitude (control) (-)		0.035				
Hansson 2000 [14], Sweden	Age		0.000				Back function (Hannover ADL) 12 months (multiple regression)
N at follow up: 455	Gender (female)		0.013				
	Psychological demand		0.024				
	Surgery during first year		0.017				

Table 3 continued

Study	Predictive factors/instruments	R^2	*P*	Δ R^2	Δ*P*	*β*	Outcome
Hansson 2000 [14], USA	Age		0.000				Back function (Hannover ADL) 12 months (multiple regression)
N at follow up: 413	Gender (female)		0.002				
	Treatment received before sick report		0.043				
	Psychological demands		0.004				
	Physical job demands		0.018				
	Back-School program		0.008				
Klenerman 1995 [23]	Physical (side flexion, neurological deficit, sagittal extension, nerve root thethering sagittal flexion and ability to sit)	0.206	0.0002				Pain and disability at 12 months
N at follow up 162	Psychosocial (inappropriate signs and symptoms, MSPQ, Roland and Morris disability, Zung depression, pain, Oswestry disability index)	0.369	0.0001				
	Fear-avoidance	0.226	0.0001				
	All	0.488	0.0001				
Truchon, 2005 [36]	Age (–), gender (female), education (–), disability	0.26		0.26	<0.01		Roland Morris at 6 months
N at follow up: 321	CPCI scales: guarding (+), resting (–), asking for assistance, relaxation, task persistence, exercise/stretch (–), coping self-statements, seeking social support	0.35		0.09	0.05		
	Catastrophizing (CSQ)	0.36		0.00	0.51		
Williams 1998 [47] Same population as [13] and	Disability at 2 months	0.32				0.37	Disability at 6 months
N at follow up: 82	Ethnicity (Minority)	0.35				–0.11	
	Orthopedic impairment	0.41				0.29	
	Job satisfaction (–)	0.48				–0.28	

Negative signs in the parentheses besides the predictors indicate that the factor is negatively related with restricted function, e.g. income (–) means that less income is related with more restricted function. If no sign is besides the predictors this means that this factor is positively related with persistent restriction in function, e.g. older age is related with more restricted function, or heat or cold therapies are related with restricted function. R2: proportion of the variance explained by the model. *P*: *P*-value, Δ R2: additional variance explained when the given factor is added to the model, and Δ*P* its corresponding *P*-value. *β*: regression coefficients

SCL-90 Somatization: Symptoms Checklist-90 Revised, Somatization subscale. SCL-90 R Depression: Symptoms Checklist-90 Revised, Depression subscale

Roland Morris: disability questionnaire with 24 items, higher scores indicates more restriction in functioning

Decision latitude: decision possibilities and control in job

Orthopaedic impairment: Waddel impairment index for measuring the severity of orthopaedic disease in back disorders. Score from 0 to 50 (more impairment)

Table 4 Studies reporting odds ratios. Predictive instruments and models used for function-related outcomes at different time-points. Models predict restricted function

Study	Predictive factors/ instruments	Odds ratio		Outcome
Karjalainen 2003 [20]	Age	3.57 (2.55–4.58)		Oswestry disability index at 12 months
N at follow up: 156	BMI	2.63 (1.84–3.42)		
	Intensity of pain at baseline	2.21 (1.54–2.89)		
	Oswestry disability index at baseline	5.22 (3.79–6.65)		
	Perceived risk of not recovering (0–10)	2.21 (1.54–2.89)		
Loisel 2002 [27]	QTFC 1 (pain without radiation)	1		Functional disability (Oswestry) at 12 months
N at follow up: 90	QTFC 2 (pain with proximal radiation (above the knee))	2.80 (0.44–17.73)		
	QTFC 3–4 (Pain w. distal radiation & Pain w. dist. radiation and neurologic signs)	3.68 (0.76 to 20.48)		
Loisel 2002 [27]	QTFC 1	1		Functional disability (SIP) at 12 months
N at follow up: 90	QTFC 2	3.15 (0.47–21.92)		
	QTFC 3–4	5.72 (1.18–36.01)		
Van der Weide 1999 [39]	Intervention group	0.92 (0.32–2.6)		Functional disability (Roland disability questionnaire) 3 months
N at follow up: 108	Avoidance coping style	4.1 (1.20–15.0)		
	Pain intensity at inclusion (OR per 10 scale units)	1.5 (1.1–2.0)		
Van der Weide 1999 [39]	Intervention group	2.2 (0.7–7.0)		Functional disability (Roland disability questionnaire) 12 months
N at follow up: 107	Lack of energy	9.9 (2.4–41.0)		
	Social isolation	4.3 (1.3–14.0)		
	Emotional effort (per 10 units)	0.6 (0.4–0.9)		Cut-off point for functional disability = 75th percentile
	Lack of variation in work (per 10 units)	1.3 (1.1–2.0)		
Von Korff 1993 [42]	Age 18–24	1.00		Poor outcome at 12 months (Grade III: high disability-moderately limiting back pain, Grade IV: high disability-severely limiting back pain)
N at follow up 999	Age 25–44	0.48	$P = 0.09$ NS	
	Age 45–64	0.67	NS	
	Age 65–74	0.70	NS	
	Male	1.00		
	Female	1.53	$P = 0.04$	Cut-off point for functional disability = 75th percentile
	College	1.00		
	graduate	1.85	$P = 0.005$	
	High School graduate	3.17	$P = 0.004$	
	<12 years			
	White	1.00		
	Black	1.21	NS	
	Other	1.02	NS	
	Pain grade I	1.00		
	Pain grade II	1.77	$P = 0.06$	
	Pain grade III	6.14	$P < 0.001$	
	Pain grade IV	8.10	$P < 0.001$	

Table 4 continued

Study	Predictive factors/ instruments	Odds ratio	Outcome
	Pain days 1–30	1.00	
	Pain days 31–89	2.45	$P = 0.002$
	Pain days 90 +	4.64	$P < 0.001$
	Recency <6 months	1.00	
	Recency 6 months to 2 years	0.95	NS
	Recentness 2–10 year	1.31	NS
	Recentness 11+year	1.99	$P = 0.02$
	Disability payments:		
	Never	1.00	
	Past only	1.51	NS
	Current	1.08	NS

Odds ratios above 1 indicate higher chance for restricted function (risk factor), values below 1 indicate reduced chance for restricted function (protective factor)

P values (reported where no confidence interval was available): whether the odds ratio for each category is significantly different from one (1) after adjusting for covariates

QTFC Quebec Task Force Classification: a commonly used classification based on simple clinical criteria including signs and symptoms (pain and neurological examination data), imaging test results, and response to treatment (in this study only QTFC categories 1–4 were used; *1* Pain without radiation, *2* Pain with proximal radiation (above knee), *3* Pain with distal radiation (below knee), *4* Pain distal radiation and neurological signs

SIP Sickness Impact Profile

Intervention group In the two analyses from van der Weide et al. patients were randomized to one of two groups: an occupational physician group (intervention group) or a reference group. Belonging to the intervention group was not statistically significant associated with more restricted function at follow up, but was retained as a control-factor

Recentness time since first back pain episode

Physical therapy, Heat and cold (treatment), Back School: patients who had physical therapy, treatment with cold and heat or attended Back-Schools had more restricted function at follow up

Table 6 shows predictive values for the dichotomized functional outcome. Diagnostic odds ratio for impaired function ranged from 4.0 to 28.7; mean 11.3 (SD 11.8). The percentages of the overall correct classified ranged from 40 to 72%; mean 55% (SD 14.65)

The sensitivity (correct classified regarding functional limitations) ranged from 63 to 91%; mean 77% (SD 13.8). Specificity (correct classified regarding good function) ranged from 29 to 93%; mean 63% (SD 26.8). Positive predictive value ranged from 22 to 98%; mean 59% (SD 39.8), negative predictive value ranged from 35 to 96% with a mean of 67% (SD 33.0).

The decision rule of Wahlgren [46] had a high diagnostic odds ratio and good overall classification for the prediction of functional outcome at 6 months, but with wide confidence intervals. Dionne and colleagues [9] evaluated generalizability. The diagnostic odds ratio was reduced from 8.27 to 4.1 because of the decreased specificity (from 57 to 29%) of the predictive tool in the new population. The overall correct classification was moderate in the development sample, but low in the new population.

Leroux [25] reported only associations between the independent variables and the one-year Roland Morris score from an analysis of covariance; these results are not shown in our tables.

Table 7 shows the large diversity of predictors included to predict functional related outcomes. Factors that are modifiable by treatments and were consistent predictors of function-related outcomes were function at baseline (measured with questionnaires) (reported 9 times), depression (8), somatization (3), psychological demand (3) and avoidance coping strategies (twice, once as avoidance coping style and once as guarding, which also concerns the avoidance of physical activities). Pain intensity was related with functional limitations positively (7) and negatively (once). The number of pain days were related positively (3), radiating pain, pain combined with disability, and the number of pain sites were each positively related with functional limitations once. The non-modifiable factors age, gender, education were important in several instruments: higher age was associated with higher functional limitations in nine populations, in one study, younger age was related with functional limitations and one study

Table 5 Studies reporting probabilities. Predictive instruments and models used for disability related outcomes at different time-points

Study	Predictive factors/instruments	Probability	Outcome
Carey 2000 [4]	4 weeks functional status: if patient reported an ability to perform his usual daily activities as well as he did before the onset of the LBP episode)	0%	probability of being functionally impaired at 3 months
N at follow up: 1246	If impaired (self declared by the patient) at 4 weeks	Percent being functionally impaired at 3/6/22 months	
		45/36/56%	
Dionne 1997 [10]	If SCL-90-R Depression score <0.444	2.0% (ts)	More than 50% on the Modified Roland-Morris score at 2 years
Decision Rule		4.9% (vs)	
N at follow up: 569 training sample, 644 validation sample	If SCL-90-R Depression score ≥ 0.444 but <1.5	1.2% (ts) 03.5% (vs)	
	and SCL-90-R Somatization <0.333		
	and SCL-90-R Somatization ≥ 0.333	19.5% (ts)	
		19.4% (vs)	
	If SCL-90-R Depression ≥ 1.5	42.4% (ts)	
		35.9% (vs)	
Dionne 2005 [9] *N* at follow up: 860	If SCL-90-R Depression score <0.444	6.2%	More than 50% on the Modified Roland-Morris score at 2 years
Decision Rule developed in Dionne 1997	If SCL-90-R Depression score ≥ 0.444 but <1.5	6.7%	
	and SCL-90-R Somatization <0.333		
	and SCL-90-R Somatization ≥ 0.333	14.2%	
	If SCL-90-R Depression ≥ 1.5	26.8%	
Von Korff 2005 [43]		Percent with Chronic pain grade II to IV at 1/2/5 years	
N at follow up	Low risk score 0–7	10.9/8.0/11.7%	
1 year: 1128	Intermediate Risk score 8–15	32.1/27.6/24.6%	
2 years: 1024	Possible chronic Pain score 16–21	58.7/51.7/46.4%	
5 years: 819	Probable chronic pain score 22 +	82.1/71.7/68.2%	

Probability of patients having different outcomes at a given follow-up time-points if predictive factor present

Risk score from Von Korff et al. 2005: risk score based on pain severity (combining with a scoring system average pain intensity, worst pain intensity and current pain intensity), and interference with pain, interference with work/housework, interference with family/social activities, and days of activity limitations due to back pain in the prior six months, and SCL-90-R-Depression score, and number of other pains, and number of days with back pain in the prior six months

Chronic pain grades in the study of Korff et al. 2005: (*1*) no back pain, (*2*) mild back pain, (*3*) moderate back pain and limitation, and (*4*) severe limiting back pain

showed a U-shaped relationship between age and function [42]. To be a woman was related with higher functional limitations in six populations. Furthermore, diverse medical interventions were related with functional limitations.

Discussion

We systematically reviewed the literature to find predictive models and tools for the transition from subacute to chronic non-specific LBP with persistent restriction in function and we analysed the methodological quality. We found instruments with limited ability to predict or explain function-related outcomes in patients with non-specific LBP. The methodological quality related to predictive issues was moderate, especially regarding the selection processes for risk factors.

Predictive tools should contain risk factors and protective factors for problems with function-related outcomes from all relevant domains (biomedical, psychosocial, occupational, social, and patient expectations about recovery of functioning). If we evaluate the predictive tools in the light of guidelines for LBP [3, 21, 40] and systematic reviews on risk factors [15, 26, 32, 37, 38] we conclude that not all known risk factors were assessed and included in the instruments. In the case of LBP and its outcome function these factors would probably be age, gender, marital status, perceived disability, pain intensity, poor expectations for recovery of function, general and specific self-efficacy, somatization, pain catastrophizing, fear

Table 6 Predicting values for dichotomized functional outcome

Study	Predictive factors	Diagnostic Odds Ratio (sens/1-sens) / (1-spec/spec)	Overall Correct classified	Correct classified functional limitation	Correct classified good function	Positive predictive value: function limitation when predicted function limitation	Negative predictive value: good function when predicted good function	Outcome	Not having good function/*N* at follow up (percentage)
Wahlgren 1997 [46]	Having more than two points on the Descriptor Differential Scale (DDS) and more than 10 points on the Beck Depression Index or the Sickness Impact Profile, compared with having less points on these scales	28.7 (3.71–1241.96)	72% (61–82)	67% (54–79)	93% (68–99.8)	98% (87–99.9)	41% (25–59)	Not to resolve in the disability score at 6 months	61/76 (80%)
Wahlgren 1997 [46]	as above	4.04 (1.11–16.36)	64% (53–75)	63% (49–75)	71% (44–90)	88% (74–96)	35% (20–54)	Not to resolve in the disability score at 12 months	59/76 (77%)
Dionne 1997 [10] validation sample	Clinical rule 1. if SCL Depression >0.444 2. if SCL Depression ≥ 0.44 but <1.5 and 2a. SCL Somatization <0.333 2b. SCL Somatization ≥ 0.333 3. if SCL Depression ≥ 1.5	8.27 (4.52–16.11)	62% (58–65)	86% (78–92)	57% (53–61)	27% (22–32)	96% (93–98)	Roland Morris ≥ 50% at 2 years	101/644 (16%)
Dionne 2005 [9]	Clinical rule developed in Dionne 1997[10]	4.1 (2.29–7.87)	40% (37–43)	91% (85–95)	29% (26–32)	22% (19–25)	94% (89–96)	Roland Morris ≥ 50% at 2 years	155/860 (18%)

Numbers in parentheses are confidence intervals, except for the last column (percentages)

Table 7 Predictors in different risk domains

	Medical variables	Self reported functioning/ disability	Occupational factors	Psychosocial factors	Sociodemographic	Management (Interventions)
Carey 2000 [4]		Good functional status at 4 weeks (–)/ self declared impairment (+)				
Dionne 1997 [10]	Number of pain days in the past 6 months (+)	Roland-Morris (+)		SCL-90-R Depression (+) SCL-90-R Somatization (+)		
Dionne et al. 1997 [10]decision rule				SCL-90-R Depression (+) SCL-90-R Somatization (+)		
Dionne 2005 [9]				SCL-90-R Depression (+) SCL-90-R Somatization (+)		
Epping-Jordan 1998 [13] Prediction for 12 months	Pain intensity (Descriptor Differential Scale) (+/-)	Disability score at 2 months (Sickness Impact Profile) (+)		Depression (Beck depression inventory) (+)	Income (–), Minority Ethnicity (+), Age (+)	
Epping-Jordan 1998 [13] Prediction for 6 months	Pain intensity (Descriptor Differential Scale) (+)	Disability score (SIP Sickness Impact Profile) (+)		Depression (+) (Beck Depression Inventory) (+)	Income (+), Minority Ethnicity (+), Age (–)	
Hansson 2000 [14] Denmark					Age (+), Gender (female +)	Physical therapy (+)
Hansson 2000 [14] Germany				Psychological demand (+)	Age (+)	Treatment (heat and cold) (+)
Hansson 2000 [14] Israel					Age (+)	
Hansson 2000 [14] Netherlands			Decision latitude (control) (–)		Age (+), Gender (female +)	
Hansson 2000 [14] Sweden	Surgery during first year (+)			Psychological demand (+)	Age (+), Gender (female +)	

Table 7 continued

	Medical variables	Self reported functioning/ disability	Occupational factors	Psychosocial factors	Sociodemographic	Management (Interventions)
Hansson 2000 [14] US			Physical job demands (+)	Psychological demand (+)	Age (+), Gender (female +)	Treatment received before sick leave (+) Back-School program (+)
Karjalainen 2003 [20]	Intensity of pain (+) BMI (+)	Oswestry disability index (+) Perceived risk of not recovering (+)			Age (+)	
Klenerman 1995 [23]	Pain (+) Side flexion, sagittal extension& flexion, neurological deficit, nerve root thethering, ability to sit	Oswestry disability index (+) Roland and Morris disability (+)		Inappropriate signs and symptoms, MSPQ, Zung Depression (+) Fear-avoidance (+)		
Loisel 2002 [27]	Radiating pain (QTFC) (+)					
Truchon 2005 [36]		Disability (+)		Coping (CPCI): Guarding (+), Resting (–), Asking for assistance (+), Relaxation (+), Task persistence (+), Exercise/Stretch (–), Coping self-statements (+), Seeking social support (+). Catastrophizing (+)	Age (–), Gender (female +), Education (–)	
Van der Weide 1999 [39] Predicting 3 months	Pain intensity (+)			Avoidance coping style		Intervention group (+)

Table 7 continued

	Medical variables	Self reported functioning/ disability	Occupational factors	Psychosocial factors	Sociodemographic	Management (Interventions)
Van der Weide 1999 [39] Predicting 12 months			Lack of variation in work (+)	Lack of energy (+) Social isolation (+) Emotional effort (–)		Intervention group (+)
Von Korff 1993 [42]	Pain Grade (+) Pain history (days with pain) (+) Recentness (–)				Age (U-shaped relationship), Gender (female +), Education (less than 12 years +), Ethnicity (Black +)	Disability payments (+)
Von Korff 2005 [43]	Pain severity (+) Pain history (days with pain) (+) Number of pain sites (+)			SCL-90 R Depression scale (+)		
Wahlgren 1997 [46]	Pain intensity (Descriptor Differential Scale) (+)	Sickness Impact Profile (+)		Depression (Beck Depression Index (+)		
Williams 1998 [47]	Orthopaedic impairment (+)	Sickness Impact Profile (+)	Job satisfaction (–)		Minority Ethnicity (+)	

Latent variables that remained in the final models and instruments. A positive sign in the parentheses indicates that this factor is positively related to restricted function, i.e. age (+) indicates that older age is associates with an increased chance of restricted function at follow up, or: treatment (Heat and Cold) (+) is a risk factor for restricted function. A negative sign in the parentheses indicates a negative relationship with restricted function, i.e. Income (–) indicates that less income is related to persistent function restriction

avoidance beliefs, distress, anxiety, health locus of control, coping strategies, symptoms of depression, pain behaviour, work intensity (heavy work and fast work pace), job tenure, availability of modified duty, perceptions of work (e.g. job satisfaction, monotonous work, job (or housework) stress, beliefs that work, housework or other activities are dangerous for the back, emotional effort concerning work or housework, social support or control at workplace) and delayed coordinated care.

Several medical interventions showed to be predictors of functional limitations. We cannot conclude whether patients with bad prognoses received more interventions, or if the interventions themselves had a causal relationship with functional limitations. One might argue that these therapies could have increased other risk factors as, for example, fear avoidance beliefs. For a discussion of the evidence of treatments in acute and subacute back pain see [40].

How to "operationalise" these constructs (e.g. distress, coping strategies, or pain behaviour) in a standardised manner should be defined to improve comparability and transferability of the predictive instruments.

Furthermore, the automatic selection processes used in the regression analyses in some of the studies included could lead to biased regression coefficients (on average too large coefficients) and to unstable variable selection, i.e. minor changes in the data may lead to selection of different predictors. For example, if a factor strongly influences the prognosis of a patient, but has a low prevalence in the population and is underrepresented in the sample in which a clinical rule is derived, then this strong predictor will not be a significant predictor in a statistical regression model and not be selected by automatic selection processes and, finally, be missing in the instrument [33]. If the instrument is later applied on a patient with this risk factor, the prognosis will be overly optimistic because the important risk factor was not assessed by the instrument. Therefore, one should not rely only on automatic selection processes. One possibility is to fit a regression model first with (a) a set of predictors that have shown to be related to the desired outcome in several studies (see list above) and with (b) a set of factors which might theoretically justify the predictive value.

Some factors, such as heavy work and fast work pace, could be summarized into a predictive index (e.g. work intensity) to reduce the number of variables in the model. Instead of relying on an automatic (stepwise) procedure, one could compare models with a "minimal" set of predictors with models including more predictors using a likelihood test to evaluate whether the models including more predictors have significant additional predictive value.

With respect to generalizability, only one study applied the model in a new population, an important step in validation: Dionne et al. [9] applied a model in a different population and obtained similar predictive values, but with decreased specificity. They considered over 100 variables in the development process and included enough patients in the development sample and the validation sample (860 patients at follow up).

Risk factors may change over time. Therefore, in studies where predictive models were devised or evaluated, analyses based on repeated measurements of predictors may be used [16]. Nevertheless, this was not done in the studies selected.

Limitations of this review

Our search in electronic databases was systematic for articles published before July 2004. We included three studies published after July 2004 (one found by expert contacts, two by searching "Related Articles" in Medline), but the search string was not rerun. Therefore, we cannot exclude that we missed some relevant studies published after July 2004.

We wanted to include studies that assessed risk factors during the period of 2–12 weeks after onset of a new back pain episode. However, differences in defining a back pain episode and the actual assessment time point could have led to an inconsistency in the selection process regarding this selection criterion. We included some studies with patients having long pain duration, for example the studies from Dionne et al. [9, 10]. Our argument to include these studies in spite of the long duration of pain is that a patient may be suffering from back pain for years without being restricted in his/her functioning. Most of the people suffering from an attack of LBP will reduce activities of daily living for some days, but will resume most activities within days. There are possibly some activities they will abandon, but a majority of people (approximately 90%) with a LBP episode are not that constrained by the back pain that they would consult a medical doctor [6]. If, however, a patient becomes restricted in activities of daily living and therefore visits a doctor, a new situation "the acute phase of an LBP episode with restricted activities of daily living" has begun. Therefore, we counted the 2–12 weeks from this time-point on (when the patient contacted a medical doctor because of LBP).

Future research

We propose the construction of a comprehensive predictive model for outcome related restriction in function. Building a model containing all relevant risk factors assembled in the many systematic reviews on risk factors for consequences of LBP will enhance information gained from such a decision tool. In a further step, the predictive values

should be evaluated in patients with LBP at about 6 weeks after the onset of a new episode leading to functional limitations; we hypothesise that at this time-point accurate prediction of prognosis will have the highest impact on clinical practice and costs. The validation process should follow expert recommendations [2, 19, 24, 41, 48]. According to Altman and Lyman [1], multiple separate and uncoordinated studies may delay the process of defining the role of prognostic markers, therefore we suggest an international coordinated study.

Implication for practice

The instruments evaluated in this review do not provide optimal information for the allocation of health care and social work resources, since there is no instrument that includes a comprehensive set of risk factors from all relevant domains of health care, psychology and social work providers. There is evidence that most of the risk factors shown in table 7 are modifiable (see e.g. [28] for pain, self-reported function, fear avoidance beliefs about physical activity, and depression, or [40] for an overview). An assessment including a comprehensive instrument would allow an informed assignment of health care, psychological and social work resources towards the modifiable risk factors and improve the triage between inexpensive standard interventions and expensive (in sense of money and time) coordinated interdisciplinary rehabilitation programs.

Clinical research, e.g. randomized trials and outcome studies, without an instrument that accurately identifies prognostic factors makes it difficult firstly to balance the prognostic factors (e.g. by stratification or minimisation), secondly, to adjust for the case mix and, thirdly, to control for confounders by multivariable analyses [31].

With the reviewed instruments, one might accurately classify patients at the extreme both ends of the "no risk–high risk" continuum, but patients in the middle "grey" zone cannot be classified accurately. As a consequence of this lack of accuracy, expensive treatments will be assigned to those patients who would even have improved with minimal interventions. On the other hand, patients who would have improved with an intensive multimodal rehabilitation program might only receive minimal interventions and pain and restrictions persist. Nevertheless, even if a clear and accurate decision cannot be made by the use of one of these instruments, using a combination of them (to address all risk domains: biological, medical, psychological, socioeconomic as well as occupational) will reduce uncertainty and provide information to apply effective interventions towards the modifiable risk factors.

Conclusions

The instruments reviewed had only limited ability to predict or explain function-related outcomes. Using one of the presented tools would provide limited information on the spectrum and amount of risk factors involved.

To provide clinicians with an accurate predictive instrument, a comprehensive predictive model should be devised by assessing known and putative risk factors (e.g. age, gender, pain intensity and history, treatments in the past, Body Mass Index, self-reported function, neurological signs, depression, somatization, fear avoidance beliefs, self-efficacy, coping strategies, physical and psychological job demands) in a sufficiently large population. The model should then be applied to different populations to assess external validity.

Acknowledgments The authors whish to thank Dr. Pius Estermann, IT specialist, for the construction of the search string and for performing the search; Leokardia Zurek, librarian, for preparing the fulltext articles; Daria Buson for assessing the Italian and Spanish articles; Leanne Pobjoy for manuscript preparation; and the experts contacted for their help. This manuscript was written as a part of a Master of Physiotherapy Sciences at the University of Maastricht and the University of Zurich. Funding: Dr. Bachmann's work was supported by the Swiss National Science Foundation (grants no. 3233B0-103182 and 3200B0-103183). Furthermore, the project was supported by the National Research Programme NRP53 of the Swiss National Science Foundation, project "Interdisciplinary intervention strategy for chronic MSDs" (module: intervention studies with groups at high risk), 2004–2009.

References

1. Altman DG, Lyman GH (1998) Methodological challenges in the evaluation of prognostic factors in breast cancer. Breast Cancer Res Treat 52:289–303
2. Altman DG, Royston P (2000) What do we mean by validating a prognostic model? Stat Med 19:453–473
3. Burton AK, Balague F, Cardon G, Eriksen HR, Henrotin Y, Lahad A, Leclerc A, Muller G, van der Beek AJ (2006) Chapter 2. European guidelines for prevention in low back pain: November 2004. Eur Spine J 15(Suppl 2):S136–S168
4. Carey TS, Garrett JM, Jackman AM (2000) Beyond the good prognosis. Examination of an inception cohort of patients with chronic low back pain. Spine 25:115–120
5. Concato J, Feinstein AR, Holford TR (1993) The risk of determining risk with multivariable models. Ann Intern Med 118:201–210
6. Croft PR, Papageorgiou AC, Thomas E, Macfarlane GJ, Silman AJ (1999) Short-term physical risk factors for new episodes of low back pain: prospective evidence from the South Manchester back pain study. Spine 24:1556–1561
7. Crook J, Milner R, Schultz IZ, Stringer B (2002) Determinants of occupational disability following a low back injury: a critical review of the literature. J Occup Rehabil 12:277–295
8. Davis KG, Heaney CA (2000) The relationship between psychosocial work characteristics and low back pain: underlying methodological issues. Clin Biomech 15:389–406

9. Dionne CE (2005) Psychological distress confirmed as predictor of long-term back-related functional limitations in primary care settings. J Clin Epidemiol 58:714–718
10. Dionne CE, Koepsell TD, Von KM, Deyo RA, Barlow WE, Checkoway H (1997) Predicting long-term functional limitations among back pain patients in primary care settings. J Clin Epidemiol 50:31–43
11. Downs SH, Black N (1998) The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health 52:377–384
12. Egger M, Smith G, Altmann D (2001) Systematic reviews in health care. Meta-analysis in context, 2nd edn. BMJ Books, London
13. Epping Jordan JE, Wahlgren DR, Williams RA, Pruitt SD, Slater MA, Patterson TL, Grant I, Webster JS, Atkinson JH (1998) Transition to chronic pain in men with low back pain: predictive relationships among pain intensity, disability, and depressive symptoms. Health Psychol 17:421–427
14. Hansson TH, Hansson EK (2000) The effects of common medical interventions on pain, back function, and work resumption in patients with chronic low back pain: A prospective 2-year cohort study in six countries. Spine 25:3055–3064
15. Hartvigsen J, Lings S, Leboeuf Yde C, Bakketeig L (2004) Psychosocial factors at work in relation to low back pain and consequences of low back pain; a systematic, critical review of prospective cohort studies. Occup Environ Med 61:e2
16. Hoogendoorn WE, Bongers PM, de Vet HC, Twisk JW, van Mechelen W, Bouter LM (2002) Comparison of two different approaches for the analysis of data from a prospective cohort study: an application to work related risk factors for low back pain. Occup Environ Med 59:459–465
17. Hopayian K (2001) The need for caution in interpreting high quality systematic reviews. Bmj 323:681–684
18. Hudak PL, Cole DC, Frank JW (1998) Perspectives on prognosis of soft tissue musculoskeletal disorders. Int J Rehabil Res 21:29–40
19. Justice AC, Covinsky KE, Berlin JA (1999) Assessing the generalizability of prognostic information. Ann Intern Med 130:515–524
20. Karjalainen K, Malmivaara A, Mutanen P, Pohjolainen T, Roine R, Hurri H (2003) Outcome determinants of subacute low back pain. Spine 28:2634–2640
21. Kendall NAS, Linton S, Main CJ (2004) New Zealand guidelines group
22. Kendall NAS, Linton SJ, Main CJ (2004) Guide to assessing psycho-social yellow flags in acute low back pain: risk factors for long-term disability and work loss. ACCatNZG Group, Wellington, New Zealand
23. Klenerman L, Slade PD, Stanley IM, Pennie B, Reilly JP, Atchison LE, Troup JD, Rose MJ (1995) The prediction of chronicity in patients with an acute attack of low back pain in a general practice setting. Spine 20:478–484
24. Laupacis A, Sekar N, Stiell IG (1997) Clinical prediction rules. A review and suggested modifications of methodological standards. Jama 277:488–494
25. Leroux I, Dionne Clermont E, Bourbonnais R (2004) Psychosocial job factors and the one-year evolution of back-related functional limitations. Scand J Work Environ Health 30:47–55
26. Linton SJ (2001) Occupational psychological factors increase the risk of back pain: a systematic review. J Occup Rehabil 11:53–66
27. Loisel P, Vachon B, Lemaire J, Durand MJ, Poitras S, Stock S, Tremblay C (2002) Discriminative and predictive validity assessment of the quebec task force classification. Spine 27:851–857
28. Mannion AF, Muntener M, Taimela S, Dvorak J (2001) Comparison of three active therapies for chronic low back pain: results of a randomized clinical trial with one-year follow-up. Rheumatology 40:772–778
29. McIntosh G, Frank J, Hogg JS, Hall H, Bombardier C (2000) Low back pain prognosis: structured review of the literature. J Occup Rehabil 10:101–115
30. Pengel LH, Herbert RD, Maher CG, Refshauge KM (2003) Acute low back pain: systematic review of its prognosis. BMJ (British Medical Journal) 327:326–330
31. Senn S (2004) Controversies concerning randomization and additivity in clinical trials. Stat Med 23:3729–3753
32. Shaw WS, Pransky G, Fitzgerald TE (2001) Early prognosis for low back disability: intervention strategies for health care providers. Disabil Rehabil 23:815–828
33. Simon R, Altman DG (1994) Statistical aspects of prognostic factor studies in oncology. Br J Cancer 69:979–985
34. Staal JB, Hlobil H, van Tulder MW, Waddell G, Burton AK, Koes BW, van Mechelen W (2003) Occupational health guidelines for the management of low back pain: an international comparison. Occup Environ Med 60:618–626
35. Teasell RW, Bombardier C (2001) Employment-related factors in chronic pain and chronic pain disability. Clin J Pain 17:S39–S45
36. Truchon M, Cote D (2005) Predictive validity of the chronic pain coping inventory in subacute low back pain. Pain 116(3):205–212
37. Truchon M, Fillion L (2000) Biopsychosocial determinants of chronic disability and low-back pain: a review. J Occup Rehabil 10:117–142
38. Valat JP, Goupille P, Vedere V (1997) Low back pain: risk factors for chronicity. Rev Rhum 64:189–194
39. van der Weide WE, Verbeek JH, Salle HJ, van Dijk FJ (1999) Prognostic factors for chronic disability from acute low-back pain in occupational health care. Scand J Work Environ Health 25:50–56
40. van Tulder M, Becker A, Bekkering T, Breen A, del Real MT, Hutchinson A, Koes B, Laerum E, Malmivaara A (2006) Chapter 3. European guidelines for the management of acute nonspecific low back pain in primary care. Eur Spine J 15(Suppl 2):S169–S191
41. Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD (2005) Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. J Clin Epidemiol 58:475–483
42. Von Korff M, Deyo RA, Cherkin D, Barlow W (1993) Back pain in primary care. Outcomes at 1 year. Spine 18:855–862
43. Von Korff M, Miglioretti DL (2005) A prognostic approach to defining chronic pain. Pain 117:304–313
44. Waddell G (2004) The back pain revolution. 2nd edn. Churchill Livingstone, Edinburgh
45. Waddell G, Burton AK, Main CJ (2003) Screening to identify people at risk of long-term incapacity for work. The Royal Society of Medicine Press Limited, London
46. Wahlgren DR, Atkinson JH, Epping JJE, Williams RA, Pruitt SD, Klapow JC, Patterson TL, Grant I, Webster JS, Slater MA (1997) One-year follow-up of first onset low back pain. Pain 73:213–221
47. Williams RA, Pruitt SD, Doctor JN, Epping-Jordan JE, Wahlgren DR, Grant I, Patterson TL, Webster JS, Slater MA, Atkinson JH (1998) The contribution of job satisfaction to the transition from acute to chronic low back pain. Arch Phys Med Rehabil 79:366–374
48. Wyatt JC, Altman DG (1995) Commentary: prognostic models: clinically usefull or quickly forgotten? BMJ (British Medical Journal) 311:1539–1541