ELSEVIER

The Spine Journal 10 (2010) 847–856

THE SPINE JOURNAL

Clinical Study

The prognosis of chronic low back pain is determined by changes in pain and disability in the initial period

Martijn W. Heymans, PT, PhD[a,b,c,*], Stef van Buuren, PhD[d,e], Dirk L. Knol, PhD[c], Johannes R. Anema, MD, PhD[a,f,g,h], Willem van Mechelen, MD, PhD[a,f,g,h], Henrica C.W. de Vet, PhD[a,c]

[a]*The EMGO Institute for Health and Care Research, VU University Medical Center, 1081 BT Amsterdam, The Netherlands*
[b]*Department of Methodology and Applied Biostatistics, VU University 1081 HV Amsterdam, Institute for Health Sciences, Amsterdam, The Netherlands*
[c]*Department of Clinical Epidemiology and Biostatistics, VU University Medical Center, 1007 MB Amsterdam, The Netherlands*
[d]*TNO Quality of Life, 2301 CE Leiden, The Netherlands*
[e]*Department of Methodology and Statistics, Faculty of Social Sciences, University of Utrecht, 3508 TC Utrecht, The Netherlands*
[f]*Body@Work, Research Center Physical Activity, Work and Health, TNO-VUmc, 1081 BT Amsterdam, The Netherlands*
[g]*Department of Public and Occupational Health, VU University Medical Center, 1081 BT Amsterdam, The Netherlands*
[h]*Research Center for Insurance Medicine, AMC UWV VUmc, 1081 BT Amsterdam, The Netherlands*

Received 28 January 2009; revised 3 May 2010; accepted 1 June 2010

Abstract

BACKGROUND CONTEXT: The recovery of patients with chronic low back pain (LBP) is slow. Furthermore, it is recently proposed that chronic LBP needs a prognostic approach to determine who will develop clinically significant back pain. Therefore, it is imperative to identify prognostic factors that are mostly seen in chronic LBP patients at an early stage. This may give clinicians tailored advice to prevent chronicity or may refer to a specific intervention.

PURPOSE: To investigate the contribution of demographic, work, clinical, and psychosocial variables, including new prognostic variables as changes in pain intensity and disability status, on the development of chronic LBP.

STUDY DESIGN/SETTING: Prospective cohort data by merging data from three randomized trials (secondary analyses).

PATIENT SAMPLE: Workers (n=628) on sick leave because of subacute nonspecific LBP.

OUTCOME MEASURES: Chronic LBP for longer than 6 months (functional measure).

METHODS: Potential prognostic variables were demographic, work, clinical, and psychosocial characteristics (self-report measures). We also included as prognostic variables a clinically relevant change in pain intensity and disability status. For the selection of variables and prognostic models, bootstrapping techniques were used in combination with multivariable logistic regression. The explained variance and discrimination were used to evaluate the clinical performance of the models.

RESULTS: The variables most strongly related to chronic LBP were as follows: no clinically relevant change in pain intensity and in disability status in the first 3 months, a higher pain intensity score at baseline, and a higher score for kinesiophobia. This prognostic model had a bootstrap-corrected explained variance of 37% and a discriminative ability (c index) of 0.80.

CONCLUSIONS: Clinical-, work-, and psychosocial-related variables contribute to the development of chronic LBP. The most promising variables are a clinically relevant decrease in pain intensity and in disability status in the first 3 months. These variables are relevant for clinicians to advise

FDA device/drug status: not applicable.

Author disclosures: none.

This study is granted by the Netherlands Organisation for Health Research and Development (ZonMW). The funding body had no role in study design; in the collection, analysis, and interpretation of data; in the writing of the manuscript; and in the decision to submit the manuscript for publication. The study protocol was approved by the medical ethics committee of the VU University Medical Center in Amsterdam.

* Corresponding author. EMGO Institute, VU University Medical Center, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands. Tel.: (31) 20-4449931; fax: (31) 20-4446775.

E-mail address: mw.heymans@vumc.nl (M.W. Heymans)

doi:10.1016/j.spinee.2010.06.005

their patients with respect to preventive measures or treatment strategies.

Keywords: Low back pain; Chronic pain; Prognosis; Prospective cohort study; Bootstrapping

Introduction

Low back pain (LBP) is the most common and expensive musculoskeletal disorders in Western countries [1]. Patients may recover spontaneously [2], and about 10% will develop chronic long-term LBP [3]. Von Korff et al. [4] showed that LBP does not have to be present all the time and that it follows an irregular pattern. The pain may be present in the background for a longer time at a lower level of pain intensity and may flare up. Flare-ups are frequently seen in chronic LBP patients and are in combination with higher levels of functional limitations, like problems with work activities, also responsible for higher pain severity. Von Korff and Miglioretti [5] also showed that the definition of chronic pain needs a prognostic approach, that is, prognostic information is needed to determine the probability of developing clinically significant back pain. Therefore, it is imperative to identify prognostic factors that are mostly seen in chronic LBP patients. If these can be measured at the early stage of LBP, clinicians may give tailored advice to prevent chronicity or may refer to a specific intervention.

Until now, a number of studies have been conducted to determine which patients are predisposed to develop chronic LBP. Evidence exists for the prognostic influence of demographic, for example, age and gender [6], and clinical LBP characteristics, for example, previous episodes of LBP and severity of pain and disability [7,8]. Work-related physical and psychosocial variables that have shown to have a relationship with LBP chronicity are heavy lifting, bending of the trunk, and job satisfaction [9]. Also, the impact of psychosocial variables, such as kinesiophobia, on the development of chronic LBP has been assumed [10,11]. Pincus et al. [12] presented in a review that pain coping and fear avoidance also have a relationship with chronic LBP but that studies that evaluated these relationships were lacking.

Several studies have demonstrated that a change in pain intensity or disability during the initial LBP episode is indicative of the development of chronic LBP [13–15]. However, two of these studies only included 96 patients [13] and 76 patients [14], and the third study [15] was conducted in a primary care setting. Therefore, the prognostic value of these ''new'' variables, a clinically relevant change in pain intensity and disability status on the development of chronic LBP, has never been tested in an occupational setting including 628 patients.

The objective of this study was to investigate the prognostic value of demographic, work, clinical, and psychosocial variables. Another objective was to introduce and investigate the prognostic value of new variables a clinically important change in pain intensity and disability status on the development of chronic LBP. We assessed which combination of indicators contributed most to this evolution.

Methods

Study design

A prospective cohort study design was used by merging secondary data from three recently completed randomized controlled trials (RCTs). The first trial (Trial 1) determined the effectiveness of a behaviorally oriented graded activity program in comparison to usual care (UC) [16]. The second trial (Trial 2) compared participative ergonomics interventions and graded activity with UC [17]. The third trial (Trial 3) compared high- and low-intensity back schools with UC [18]. Trial 1's recruitment period (Staal et al. [16]) was from 1999 to 2002, and for Trials 2 and 3 (Anema et al. [16] and Heymans et al. [18]), it was from 2000 to 2003. In the Staal et al. study, 150 workers were eligible and 16 excluded (134 included). In the Anema/Steenstra et al. trial, 243 workers were eligible and 196 were randomized (one worker's information was missing in the present study), leaving 195 workers for the present study. In the Heymans et al. trial, 814 workers were eligible, 515 excluded, and 299 randomized. These RCTs were similar in design in that patients visited their occupational physician (OP) at one of the participating occupational health services when they were on sick leave for not more than 8 weeks, and the same inclusion and exclusion criteria were used. All studies had a follow-up of 1 year. The Staal et al. trial included Dutch airline company workers who were, among others, working on the luggage handling department. The Anema/Steenstra et al. trial included patients from the industrial and healthcare setting. The Heymans et al. trial included, among others, heavy steel construction workers and office workers. This means that workers from various company settings were included but that there was a shift to include more heavy duties.

Companies and occupational health services were mostly situated in the middle and western part of the Netherlands.

Data merging process

The data merging procedures were done systematically by one researcher (MWH). First, the data set of each RCT was explored to scan if all relevant prognostic variables were present. After that, the distribution and categories of these variables were checked, and variables were defined in the same way. In the next step, the data sets were merged. Patients kept their original record number, and their data remained confidential. Then, the treatment group variable was redefined so that patients within the same intervention group were assigned the same code (eg, all patients who received UC were assigned the same category number). In

the final step, we checked all variables for outliers in the data or unrealistic values.

Study population

The population consisted of blue- and white-collar workers covering a broad range of professions. Patients were eligible for participation if they met the following inclusion criteria: nonspecific LBP, defined as pain localized in the lower back without a specific underlying cause (eg, herniated disc, neurological disorder, rheumatological disorder, malignancy, infection, or fracture); sick-listed because of LBP (completely or partially) for not more than 8 weeks; between 18 and 65 years of age; and ability to complete questionnaires written in Dutch. Exclusion criteria were specific cause of the LBP, pregnancy, serious psychiatric disorders, and juridical conflict at work. Patients who met the eligibility criteria and who were willing to participate signed an informed consent form. All three RCTs were approved by the medical ethics committee of the VU University Medical Center.

Outcome measure

Pain intensity was assessed on a Numerical Rating Scale (NRS) [19] at baseline and after 3 and 6 months. The outcome measure chronic LBP was defined as having a pain intensity score of ≥4 at baseline and ≥3 at 3 and 6 months of follow-up [20]. Patients were labeled according to this definition as having chronic LBP (indicated by 1) or not (indicated by 0).

Selection of prognostic variables

The selection of relevant prognostic variables was performed in two steps. First, the literature on prognosis for chronic LBP was reviewed [21–24]. On the basis of this review and the prognostic variables in our study, we composed a list of potential prognostic variables. Second, we presented this list to 42 OPs working in different sectors. They were asked to judge, according to their expert opinion, whether each variable would contribute to the development of chronic LBP. The variables smoking (1="yes," 0="no"), body height and body weight (body mass index), years working in current job, and the level of education (low: primary education; moderate: secondary school with/without postsecondary education; and high: university) were not considered relevant after performing the literature review. This was in line with the judgment of the physicians after the e-mail round, and so we excluded these variables from our analyses.

Potential prognostic variables

The following variables were considered as candidates: age, gender (1="male," 0="female"), duration of complaints before randomization (in weeks), radiation to one or both legs (1="yes," 0="no"), and treatment during study enrollment (1="yes," 0="no"). Physical activity was measured with the Baecke questionnaire [25]. This questionnaire consists of three sections: work activity, sports activity, and nonsports leisure activity. Each section consists of questions that have to be scored on a 5-point Likert scale, ranging from never to always or very often. Final scores depend on patients' level of occupation and sport intensity. The section scores are added to obtain the final score. Pain intensity and disability status at baseline were scored by using an NRS (range 0–10) [19] and the Roland Disability Questionnaire (RDQ, range 0–24) [26], respectively. We constructed the indicators for a change in pain intensity and a change in disability status as follows. We calculated the patient's change scores by subtracting their NRS scores at 3 months from those at study inclusion. Patients were labeled "1" if they reported a reduction of at least 3 points on the NRS scale or were labeled "0" if they reported a change score of <3 points on the NRS scale. An NRS scale change of 3 points is considered a clinically relevant change score [27]. For disability status, we used a decrease of at least 4 points in the RDQ scores [27]. Potential job-related physical variables were measured by the Dutch Musculoskeletal Questionnaire [28]. The physical variables examined were daily lifting and bending and twisting of the trunk. Potential work-related psychosocial variables were measured by a Dutch version of the Job Content Questionnaire. For the Job Content Questionnaire, patients score certain aspects of their work, for example, if they have to work fast, if they learn new things from their work, and so forth, on a 4-point scale with the answer categories: "totally disagree," "disagree," "agree," and "totally agree." These scores are combined into the

EVIDENCE & METHODS

Context
Determining which patients are at risk for developing chronic low back pain may be important to facilitate active early intervention and, hopefully, recovery.

Contribution
The authors found that patients who reported severe back pain and who reported their pain has not improved over 3 months have abnormal kinesiophobia scores and were more likely to have continued low back pain lasting longer than 6 months.

Implication
These findings corroborate previous reports implicating reported pain intensity and psychometric findings with persistent back pain disorders. However, the questions remain: Are these variables causative? Why do some patients report greater pain intensity than others? And, are these findings a function of local tissue damage or central pain processing events?

—*The Editors*

following dimensions: job control (consists of quantitative job demands and decision authority), job demands (consists of skill discretion), and social support (consists of supervisor support and coworker support) [29]. Job satisfaction was assessed by means of a question concerning job task satisfaction consisting of the answer categories: ''poor,'' ''reasonable,'' ''moderate,'' and ''good.'' The Dutch version of the Tampa Scale for Kinesiophobia was used to measure the extent to which people feared that exercise can lead to reinjury (range 17–68) [30]. A high score indicates much fear for physical activity or injury. Fear of movement, avoidance of activities, and back pain beliefs were measured with the Fear Avoidance Beliefs Questionnaire work scale (range 0–42) [31]. Coping with pain was measured with the Pain Coping Inventory questionnaire, which measures cognitive and behavioral coping strategies. The questionnaire consists of six subscales that are divided into ''active'' and ''passive'' pain coping strategies. Active pain coping consists of the subscales transformation of pain, distraction, and lowering demands, and passive pain coping consists of the subscales withdrawal, worrying, and resting [32]. Each subscale is rated on a 4-point scale with the following score categories: ''seldom or never,'' ''sometimes,'' ''often,'' and ''very often.'' The potential prognostic variables were assessed by means of self-reported questionnaires.

Analyses

Model building process

Logistic regression was used to examine the prognostic value of each variable. In each regression model, the outcome chronic LBP was the dependent variable and the prognostic variables were the independent variables. We adjusted for the effects of the interventions in each logistic regression model. Restricted cubic spline functions and spline plots were used to examine the potential nonlinear behavior of the continuous indicators with the outcome [33]. The use of restricted cubic spline functions allowed us to relax the linearity assumption of the logistic regression model. We did not find a nonlinear relation for any continuous indicator and, therefore, did not include spline functions in the fitted models.

Missing data

The reason that some variables had high percentages of missing values was that not all questionnaires were used in all RCTs. Trial 1 (study of Staal et al.) did not use the Job Content Questionnaire and the Dutch Musculoskeletal Questionnaire and the question about treatment during study enrollment. The Baecke questionnaire and the question on duration of complaints were not used in Trial 2 (study of Anema et al.). The Fear Avoidance Beliefs Questionnaire was not assessed in Trial 3 (study of Heymans et al.). This led to a missing data problem in these variables when the data of the three RCTs were merged. To fill in variables with missing values, we applied multiple imputation by using the Multiple Imputation by Chained Equations package [34]. This is a flexible imputation method that uses a series of regression imputation models based on the information of other variables to substitute each missing value several times. Logistic regression is used to impute incomplete dichotomous variables and linear regression to impute continuous variables. We generated 10 multiple imputed data sets.

Variable selection

Before we started the variable selection process, we checked if variables were highly correlated, that is, if colinearity existed in our models. The correlation coefficients of variables in our data set did not exceed 0.5. We, therefore, did not expect selection problems of variables because of colinearity in our models [33]. Variable and model selection in these 10 data sets were performed by a two-step bootstrap modeling approach.

Step 1. During the first step of this procedure, backward logistic regression analyses (p value of .157) were applied on 200 bootstrap samples [35]. The bootstrap samples were drawn with replacement from the imputed data sets and were of equal sample size as this original sample. On the basis of the selection frequencies in these bootstrap samples, prognostic factors were included into a second modeling step. We included factors that were selected in more than 50% of the regression models for further analyses. As a sensitivity analysis, we also included factors that were selected in more than 40% of the models [36].

Step 2. In the second bootstrap modeling step (500 bootstrap samples), again backward regression analyses were performed on each of the bootstrap samples with the same selection criterion as in Step 1. Now, the frequency of the selected models was calculated by using the factors selected in the first step. The idea is that stable models will be selected more often than unstable models. The strength of this procedure is that the final ''best'' model is chosen among several other competing models.

Final model

The logistic regression coefficients and standard errors of the final model were pooled and estimated according to Rubin's [37] rules, by averaging over the 10 imputed data sets. Subsequently, estimates were converted to odds ratios (ORs) and corresponding 95% confidence intervals. The final model was also tested for relevant interactions.

To ensure sufficient power during the modeling process, we also considered the balance between the number of variables and outcome events/nonevents (which one is lowest), that is, the events per variable, in the models [38].

Performance of the prognostic model

The performance of the prognostic model was studied in terms of discrimination and explained variance. Discrimination expresses how well the prognostic model is able to

distinguish between patients with and without chronic LBP. Discrimination was obtained by the c index, which equals the area under the receiver operating characteristic curve in logistic regression [39]. The explained variance of the prognostic models was calculated as Nagelkerke's R^2 [40]. The explained variance gives an indication of how much of the variance in the outcome can be explained by the prognostic variables. Prognostic models usually perform better in the patient data that were used to build the model than in new patient data sets because of optimism in regression coefficients and performance measures [41]. To correct for these optimistic estimates of the c index and explained variance in our data set, bootstrapping techniques were used [38]. The model performance indices were calculated on each imputed data set and averaged over the 10 imputed data sets.

Software

Imputation was applied by using Multiple Imputation by Chained Equations, and the backward selection procedures were performed with R software. We used adapted versions of the Multiple Imputation by Chained Equations and design libraries [33,34].

Results

Table 1 gives an overview of the mean values and proportions of the potential prognostic variables in the study population and of the percentage of missing values. Most of the missings occurred because some variables were included

Table 1
Patient characteristics at baseline (n=628), the observed score ranges, and the percentage of missing information for the potential prognostic variables of chronic low back pain

Characteristics	Value	Observed range	Missing (%)
Age (mean years±SD)	40.6 (9.5)	17–65	0
Gender (male, %)	71.0	—	0
Physical activity (mean±SD)	8.8 (1.0)	6–11.5	44.6
Smoking (%)	47.3	—	7.5
Body mass index (mean±SD)	25.0 (4.0)	17.5–48.1	4.0
Level of education (%)			26.1
Low: primary education	46.3	—	
Moderate: secondary school with or without postsecondary education	44.8	—	
High: university	8.8	—	
Job satisfaction (%)		—	2.8
Poor	3.5	—	
Moderate	10.4	—	
Reasonable	40.3	—	
Good	43.0	—	
Job Content Questionnaire (mean±SD)			
Job control	56.2 (9.2)	20–44	25.6
Job demands	33.1 (4.8)	24–78	24.7
Social support	22.5 (4.1)	8–32	24.8
Daily exposed to:			
>25 kg at least once a day (%)		—	24.2
Never	29.6	—	
Sometimes	19.7	—	
Quite frequently	14.8	—	
Very frequently	11.6	—	
Bending and twisting of the trunk (%)		—	24.2
Never	18.8	—	
Sometimes	18.6	—	
Quite frequently	23.1	—	
Very frequently	15.3	—	
Duration of complaints (weeks) before randomization, median (IQR)	5.8 (13.3)	2–780	33.3
Pain radiation in one or both legs (%)	33.8	—	2.1
Functional disability at baseline (RDQ) (mean±SD)	11.3 (5.2)	0–24	5.1
Treatment during study enrollment (%)	58.6	—	23.6
Pain intensity (NRS) (mean±SD)	6.2 (1.9)	0–10	3.0
Pain coping, active (mean±SD)	6.7 (1.2)	3.5–11.6	5.6
Pain coping, passive (mean±SD)	6.5 (1.3)	3.4–10.5	7.0
Fear avoidance beliefs (mean±SD)	19.5 (9.7)	0–42	48.1
Kinesiophobia (mean±SD)	39.8 (6.7)	19–62	6.2
Change in pain intensity (%)	35.8	—	19.7
Change in functional disability (%)	36.3	—	23.4

SD, standard deviation; IQR, interquartile range; RDQ, Roland Disability Questionnaire; NRS, Numerical Rating Scale.

in only two of the three RCTs. Two variables, physical activity and fear avoidance beliefs, contained the highest percentages of missing data (44.6% and 48.1%, respectively). Other variables had missing values within the range of 0% to 33.3%. After 6 months, 204 patients (33%) had chronic LBP. On the basis of the 204 patients with chronic LBP in our study and the 31 variables in the initial model, the events per variable in our models was 204/31=6.6.

The patients in the three trials were similar on most characteristics. The median time of sick leave for the patients in all the three studies was 28 days. The mean pain intensity measured on an NRS was 6.6 for the patients in Trials 1 and 3 and 5.5 in Trial 2. Patients in Trials 1, 2, and 3 had mean scores for kinesiophobia, which were measured with the Tampa scale, of 40.1, 39.0, and 39.8, respectively. With respect to disability status and gender, small differences between the three study populations existed. For disability status, measured with the RDQ, the mean scores for patients were 8.6, 14.2, and 13.2 for Trials 1, 2, and 3, respectively. The percentage of men was larger in Trials 1 and 3, with 78% and 94%, respectively, compared with 44% in Trial 2. These differences may be because of the different companies that were involved in each trial.

Table 2 shows the selection frequencies of the prognostic variables and models as a result of the bootstrap selection procedures. The selection frequencies of the prognostic variables in the 200 bootstrap samples ranged from 22.4% to 100%. The variables change in pain intensity and pain intensity at baseline had a selection frequency of 100%, that is, they were included in all prognostic models. The variable change in disability status also had a high selection frequency of 97.7%. There were seven prognostic variables: change in pain intensity, pain intensity at baseline, change in disability status, treatment, kinesiophobia, lifting, and job satisfaction, with a selection frequency in excess of 40%.

The 14 remaining prognostic variables were excluded. With the subset of seven variables, 31 different prognostic models were found in 10×500 bootstrap samples. Table 2 presents the 10 most popular prognostic models. The most popular prognostic models were found 16.2% of the time. This is the most stable model, consisting of five variables: change in pain intensity, pain intensity at baseline, change in disability status, treatment, and kinesiophobia. Using a stricter selection frequency of 50%, the same prognostic model was most popular. The popularity of the model was 30%, among 16 different models.

It is clear that models vary considerably in the number of prognostic variables and in the composition of the prognostic models. Only three prognostic variables, change in pain intensity, pain intensity at baseline, and change in disability status, were included in all prognostic models.

Table 2
Selection frequencies of variables at Step 1 and models at Step 2 as a result of the bootstrap model selection process

		Prognostic models selected in Step 2 (the first 10 models are shown)*									
Prognostic variables	Step 1 (%)	1 (5)	2 (6)	3 (6)	4 (7)	5 (5)	6 (6)	7 (4)	8 (5)	9 (4)	10 (5)
1. Change in pain intensity	100	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2. Pain intensity at baseline	100	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
3. Change in functional disability	97.1	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
4. Treatment	77.4	✓	✓	✓	✓	✓	✓	✓	✓	—	—
5. Kinesiophobia	56.2	✓	✓	✓	✓	—	—	—	—	✓	✓
6. Lifting	53.7	—	✓	—	✓	✓	✓	—	—	—	✓
7. Job satisfaction	46.2	—	—	✓	✓	—	✓	—	✓	—	—
8. Bending and twisting of the trunk	36.8										
9. Duration of complaints	36.3										
10. Physical activity	32.5										
11. Fear avoidance	31.4										
12. Job control	27.3										
13. Job demands	26.8										
14. Functional status at baseline	26.6										
15. Age	25.6										
16. Treatment during study enrollment	25.4										
17. Gender	25.1										
18. Passive pain coping	24.2										
19. Social support	23.8										
20. Active pain coping	23.5										
21. Pain radiation	22.4										
Model selection frequency†		16.2	13.7	13.4	11.6	7.1	6.8	6.4	6.0	3.4	3.2
Model selection frequency‡		30.0	25.8	§	§	13.4	§	11.6	§	5.5	5.8

✓, variables included in the model; —, variables not selected.

* The number of variables in the model is given within parentheses.

† Model contained the subset of variables with a selection frequency of >40%.

‡ Model contained the subset of variables with a selection frequency of >50%.

§ Model contained variables with a selection frequency of <50% and cannot be included in this step.

Table 3 shows the regression coefficients, 95% confidence intervals, explained variance, and c index of discrimination of the most popular prognostic model. This prognostic model shows that chronic LBP at 6 months is associated with patients who do not show a clinically relevant change in pain intensity and in disability status within the first 3 months, who have a high pain intensity score at baseline, and who have a higher score for kinesiophobia.

This prognostic model, with four variables, had a bootstrap-corrected c index of 0.80 and an explained variance of 37%. We also explored the performance of the other models in Table 2 and found out that the ORs and performance parameters were similar among prognostic models (data not shown).

Discussion

We found the following factors as predictors for the transition from subacute to chronic LBP: not experiencing a clinically relevant change in pain intensity and in disability status in the first 3 months and a higher score for pain intensity and kinesiophobia at baseline. This prognostic model had an explained variance of 37% and a discriminative ability (c index) of 0.80.

Comparison with the findings in the literature

A relevant improvement in pain intensity and disability status during the initial LBP episode as a predictor for a lower probability to develop chronic LBP has been confirmed by other studies. In all these studies, different variables to assess these changes were used. We used a change of 3 and 4 points on the NRS and RDQ, respectively, within the first 3 months after the LBP episode started. These measures are considered clinically relevant [33]. Carey et al. [13] assessed the failure to functionally improve, according to the patients' own perception, measured at 3 months after the first consultation. Dunn and Croft [15] classified patients according to their absolute scores of 5 and 14 points on pain and disability assessed with the NRS and RDQ, respectively, at the start of the study and during reassessment after 1 month. These studies imply that information that is repeatedly assessed during the initial LBP episode is relevant for distinguishing between patients who either will or will not develop chronic LBP. A new research priority that can be derived from our study is to investigate at which time point of reassessment by the physician, for example, after 1 or 2 weeks, prognostic information is already of value for the clinical setting and which pain and disability parameters are most predictive.

Table 3
The ORs and CIs of variables included in the model with the highest selection frequency

Model*	OR† (95% CI)	
Change in pain intensity‡	0.19 (0.09–0.41)	
Change in functional disability‡	0.26 (0.12–0.56)	
Pain intensity at baseline§	1.72 (1.50–1.97)	
Kinesiophobia§	1.03 (1.00–1.06)	
	Apparent	Bootstrap corrected
Discrimination (c index)	0.82	0.80
Explained variance (R^2), %	38	37

OR, odds ratio; CI, confidence interval; LBP, low back pain.

* Model is adjusted for treatment effect.

† OR>1 indicates a higher risk of chronic LBP and OR<1 a lower risk of chronic LBP.

‡ The risk of chronic LBP is 0.19 and 0.26 times lower for a patient who, respectively, experiences a clinically relevant decrease in pain intensity and functional disability compared with a patient who does not experience that change.

§ The risk of chronic LBP is 1.72 and 1.03 times higher per point increase in, respectively, pain intensity and kinesiophobia.

Consistent with the findings in our study, a higher level of pain intensity at baseline has been identified as a relevant prognostic variable in other studies before [6,8]. Pain intensity is, in combination with functional limitations or sick leave, an indication of higher severity of the LBP [42]. Patients with more severe LBP during the initial LBP episode may be less inclined to recover from their LBP and consequently are more likely to become chronic LBP sufferers [4].

According to our study, a higher score for kinesiophobia increases the risk of developing chronic LBP. The impact of kinesiophobia on the transition from subacute to chronic LBP was also reported in the studies of Heneweer et al. [43] and Swinkels-Meewisse et al. [44].

In the prognostic models in Table 2, the variables heavy lifting and job satisfaction were not always retained. This is consistent with the findings in the literature with respect to relevance of these variables. There is strong evidence in the literature that frequent exposure to heavy lifting increases the risk of developing LBP [45,46]. However, the role of this variable on the prognosis of chronic LBP is less obvious. That the patients in our study who were daily exposed to heavy lifting develop chronic LBP is an indication of the need to implement this variable in future prognostic studies. There is also evidence that psychosocial work characteristics may be relevant variables for LBP patients to eventually maintain or quit their work [47]. Williams et al. [48] showed that poor job satisfaction is an important indicator of developing chronic LBP. Therefore, job satisfaction can also be considered a relevant work-related variable in the transfer from subacute to chronic LBP.

Variables that had an effect on the prognosis of chronic LBP in previous studies [6,12] but that were not confirmed in our study were fear avoidance, disability status at baseline, age, gender, pain coping, and pain radiation.

We collected data on the most relevant potential prognostic variables for chronic LBP. The bootstrap-corrected explained variance of the prognostic model was 37%. Bekkering et al. [49] reported an explained variance of 25% for prognostic models that investigated long-term LBP. Dionne et al. [50] presented an explained variance of 30% for studying long-term functional limitations in LBP patients.

In general, these values of explained variance are not high. Furthermore, prognostic models will perform better in patient samples used to develop the model than in other samples. A consequence is that when the prognostic model is applied in another patient population, for example, patients who visit a medical specialist because they have more severe LBP, the generalizability will be less optimal. This also accounts for the explained variance, which will be lowered. The explained variance of our study indicates that moderate beneficial effects for the prevention of chronic LBP by interventions aimed at targeting the variables in our prognostic model can be expected.

Remarks on our study: definition of chronic LBP

We defined chronic LBP as a persistent pain intensity score of ≥4 at baseline and >3 on the NRS assessed at 3 and 6 months of the follow-up period. How to define chronic LBP has been considered in several studies and is an important issue, but the definition is still not standardized [4]. In a recent article, it was emphasized that chronic LBP is constantly changing and that prognostic information is needed to determine the course of LBP [4]. Dunn et al. [20] studied the pain pathways in LBP patients over a 12-month period. They concluded that there are four different pain patterns, of which three may result in chronic LBP, either with or without variations in the levels of pain [20].

A limitation of our study might be that our definition of chronic LBP is not applicable to all types of chronic LBP patients in practice. For example, patients who intermittently experience pain-free periods were not identified by our definition. Their prognosis may be determined by other prognostic variables for chronic LBP. Furthermore, our definition of chronic LBP was determined on the level of pain intensity. It has been argued that chronic LBP may be based not only on pain intensity levels but also on limitations in disability [42]. However, recent studies showed that LBP pathways are linked to the levels of functional disability in such a way that if one knows the level of pain intensity, the level of functional disability can be determined and vice versa [20].

Remarks on our study: choices in the data analysis

In our prognostic models, we adjusted for the treatment effects that were examined in each RCT that delivered the patient data. Therefore, the treatment effects may have influenced the prognosis of chronic LBP. However, the treatment effects of UC [16–18] and low- and high-intensity back schools on pain intensity were small [18]. The graded activity and a workplace intervention did not affect pain intensity [16,17]. These interventions will, therefore, have had a minor impact on the prognosis of chronic LBP. Therefore, we choose to present our final prognostic model without the treatment variable.

We have also investigated if significant interaction effects could be identified in our final model. Significant interaction effects were present for the variables pain intensity at baseline and a change in pain intensity, and pain intensity at baseline and a change in disability status. Other significant interaction effects could not be identified. However, including the significant interaction terms in our final model caused a strong unreliable increase in the ORs of the variables a change in pain intensity and a change in disability status. Probably this was because of a subgroup of patients on these variables with no or rare cases with positive scores on the outcome when the interaction terms are included to the model. This phenomenon that has also been described as ''separation'' [51] is a drawback of the Wald test statistic in logistic regression and results in the estimation of ORs that are highly biased away from 1. We, therefore, decided to omit the interaction terms from our final model.

Not all the same prognostic variables were assessed in each RCT. Consequently, after the data were merged, the percentage of missing data in some variables was around 45%. For these variables, multiple imputation produces valid results under the missing at random assumption, that is, missings in these variables can be explained by the available data in the data set. This latter statement holds in our study because there was no specific reason in each of the original trials to exclude prognostic variables (this means that for each separate trial, there is no specific reason that information was not assessed). Furthermore, most variables in our study could be used in the imputation model to estimate the missing values because they have shown to be important in LBP prognosis.

Generalizability of study results

The study population was recruited by a large group of OPs attending a broad range of different workplace settings, which enhances the generalizability of our findings. Furthermore, we were able to include almost all variables that are mentioned in the literature as prognostic factors. Moreover, the selection of relevant prognostic variables was based not only on evidence in the literature but also on clinical expertise of OPs.

Clinical implications

Our study findings are important for clinicians who are confronted with patients on sick leave because of subacute LBP. Changes in pain intensity and in disability status in the first 3 months are good predictors of LBP persistence. This means that for daily practice, physicians have to see their patients more often during the initial LBP episode to make more reliable predictions of the prognosis. If pain persists and it concerns patients with initial severe LBP who show kinesiophobia, a chronic problem may be anticipated.

Conclusions

Our study showed the importance of the variables not experiencing a clinically relevant change in pain intensity and in disability status in the first 3 months and a higher score for pain intensity and kinesiophobia at baseline on the prognosis of chronic LBP. A model with these variables showed a good explained variance and discrimination. Especially, the inclusion of the ''new'' prognostic variables that assessed a clinically relevant change in pain intensity and disability within the first 3 months may have important clinical implications.

Acknowledgments

We thank Ivan Steenstra, Bart Staal, and Hynek Hlobil for providing their data.

References

[1] Andersson GB. Epidemiological features of chronic low-back pain. Lancet 1999;354:581–5.
[2] Henschke N, Maher CG, Refshauge KM, et al. Prognosis in patients with recent onset low back pain in Australian primary care: inception cohort study. BMJ 2008;337:a171.
[3] Frank JW, Brooker AS, DeMaio SE, et al. Disability resulting from occupational low back pain. Part II: what do we know about secondary prevention? A review of the scientific evidence on prevention after disability begins. Spine 1996;21:2918–29.
[4] Von Korff M, Deyo RA, Cherkin D, et al. Back pain in primary care. Outcomes at 1 year. Spine 1993;18:855–62.
[5] Von Korff M, Miglioretti DL. A prognostic approach to defining chronic pain. Pain 2005;117:304–13.
[6] Thomas E, Silman AJ, Croft PR, et al. Predicting who develops chronic low back pain in primary care: a prospective study. BMJ 1999;318:1662–7.
[7] Coste J, Delecoeuillerie G, Cohen de Lara A, et al. Clinical course and prognostic factors in acute low back pain: an inception cohort study in primary care practice. BMJ 1994;308:577–80.
[8] van den Hoogen HJ, Koes BW, Deville W, et al. The prognosis of low back pain in general practice. Spine 1997;22:1515–21.
[9] Cedraschi C, Allaz AF. How to identify patients with a poor prognosis in daily clinical practice. Best Pract Res Clin Rheumatol 2005;19: 577–91.
[10] Kendall NA. Psychosocial approaches to the prevention of chronic pain: the low back paradigm. Baillieres Best Pract Res Clin Rheumatol 1999;13:545–54.
[11] van der Windt DA, Kuijpers T, Jellema P, et al. Do psychological factors predict outcome in both low-back pain and shoulder pain? Ann Rheum Dis 2007;66:313–9.
[12] Pincus T, Vlaeyen JW, Kendall NA, et al. Cognitive-behavioral therapy and psychosocial factors in low back pain: directions for the future. Spine 2002;27:E133–8.
[13] Carey TS, Garrett JM, Jackman AM. Beyond the good prognosis. Examination of an inception cohort of patients with chronic low back pain. Spine 2000;25:115–20.
[14] Wahlgren DR, Atkinson JH, Epping-Jordan JE, et al. One-year follow-up of first onset low back pain. Pain 1997;73:213–21.
[15] Dunn KM, Croft PR. Repeat assessment improves the prediction of prognosis in patients with low back pain in primary care. Pain 2006;126:10–5.
[16] Staal JB, Hlobil H, Twisk JW, et al. Graded activity for low back pain in occupational health care: a randomized, controlled trial. Ann Intern Med 2004;140:77–84.
[17] Anema JR, Steenstra IA, Bongers PM, et al. Multidisciplinary rehabilitation for subacute low back pain: graded activity or workplace intervention or both? A randomized controlled trial. Spine 2007;32: 291–8.
[18] Heymans MW, de Vet HC, Bongers PM, et al. The effectiveness of high-intensity versus low-intensity back schools in an occupational setting: a pragmatic randomized controlled trial. Spine 2006;31: 1075–82.
[19] Carlsson AM. Assessment of chronic pain. I. Aspects of the reliability and validity of the visual analogue scale. Pain 1983;16: 87–101.
[20] Dunn KM, Jordan K, Croft PR. Characterizing the course of low back pain: a latent class analysis. Am J Epidemiol 2006;163:754–61.
[21] van TM, Koes B, Bombardier C. Low back pain. Best Pract Res Clin Rheumatol 2002;16:761–75.
[22] Diamond S, Borenstein D. Chronic low back pain in a working-age adult. Best Pract Res Clin Rheumatol 2006;20:707–20.
[23] Crook J, Milner R, Schultz IZ, et al. Determinants of occupational disability following a low back injury: a critical review of the literature. J Occup Rehabil 2002;12:277–95.
[24] Hartvigsen J, Lings S, Leboeuf-Yde C, et al. 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 2004;61:e2.
[25] Baecke JA, Burema J, Frijters JE. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. Am J Clin Nutr 1982;36:936–42.
[26] Brouwer S, Kuijer W, Dijkstra PU, et al. Reliability and stability of the Roland Morris Disability Questionnaire: intra class correlation and limits of agreement. Disabil Rehabil 2004;26:162–5.
[27] Ostelo RW, de Vet HC. Clinically important outcomes in low back pain. Best Pract Res Clin Rheumatol 2005;19:593–607.
[28] Hildebrandt VH, Bongers PM, van Dijk FJ, et al. Dutch Musculoskeletal Questionnaire: description and basic qualities. Ergonomics 2001;44:1038–55.
[29] Karasek R, Brisson C, Kawakami N, et al. The Job Content Questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. J Occup Health Psychol 1998;3:322–55.
[30] Swinkels-Meewisse EJ, Swinkels RA, Verbeek AL, et al. Psychometric properties of the Tampa Scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. Man Ther 2003;8:29–36.
[31] Waddell G, Newton M, Henderson I, et al. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain 1993;52:157–68.
[32] Kraaimaat FW, Evers AW. Pain-coping strategies in chronic pain patients: psychometric characteristics of the pain-coping inventory (PCI). Int J Behav Med 2003;10:343–63.
[33] Harrell FE Jr. Regression modelling strategies. New York, NY: Springer, 2001.
[34] Van Buuren S, Oudshoorn K. Flexible multivariate imputation by MICE. Leiden, The Netherlands: TNO Quality of Life, 1999.
[35] Akaike H. Information theory and an extension of the maximum likelihood principle. In: Second International Symposium on Information Theory; 1973; Tsahkadsor, Armenian SSR: 267–81.
[36] Augustin NH, Sauerbrei W, Schumacher M. The practical utility of incorporating model selection uncertainty into prognostic models for survival data. Stat Model 2005;5:95–118.
[37] Rubin DB. Multiple imputation for nonresponse in surveys. New York, NY: John Wiley & Sons, 1987.
[38] Harrell FE Jr, Lee KL, Matchar DB, et al. Regression models for prognostic prediction: advantages, problems, and suggested solutions. Cancer Treat Rep 1985;69:1071–7.

[39] Vergouwe Y, Steyerberg EW, Eijkemans MJ, et al. Validity of prognostic models: when is a model clinically useful? Semin Urol Oncol 2002;20:96–107.
[40] Nagelkerke NJ. A note on a general definition of the coefficient of determination. Biometrika 1991;78:691.
[41] Vergouwe Y, Steyerberg EW, Eijkemans MJ, et al. Validity of prognostic models: when is a model clinically useful? Semin Urol Oncol 2002;20:96–107.
[42] Von Korff M, Ormel J, Keefe FJ, et al. Grading the severity of chronic pain. Pain 1992;50:133–49.
[43] Heneweer H, Aufdemkampe G, van Tulder MW, et al. Psychosocial variables in patients with (sub)acute low back pain: an inception cohort in primary care physical therapy in The Netherlands. Spine 2007;32:586–92.
[44] Swinkels-Meewisse IE, Roelofs J, Schouten EG, et al. Fear of movement/(re)injury predicting chronic disabling low back pain: a prospective inception cohort study. Spine 2006;31:658–64.
[45] Burdorf A, Sorock G. Positive and negative evidence of risk factors for back disorders. Scand J Work Environ Health 1997;23:243–56.
[46] Hoogendoorn WE, Bongers PM, de Vet HC, et al. Flexion and rotation of the trunk and lifting at work are risk factors for low back pain: results of a prospective cohort study. Spine 2000;25:3087–92.
[47] Hoogendoorn WE, Bongers PM, de Vet HC, et al. Psychosocial work characteristics and psychological strain in relation to low-back pain. Scand J Work Environ Health 2001;27:258–67.
[48] Williams RA, Pruitt SD, Doctor JN, et al. The contribution of job satisfaction to the transition from acute to chronic low back pain. Arch Phys Med Rehabil 1998;79:366–74.
[49] Bekkering GE, Hendriks HJ, van Tulder MW, et al. Prognostic factors for low back pain in patients referred for physiotherapy: comparing outcomes and varying modeling techniques. Spine 2005;30: 1881–6.
[50] Dionne CE, Koepsell TD, Von Korff M, et al. Predicting long-term functional limitations among back pain patients in primary care settings. J Clin Epidemiol 1997;50:31–43.
[51] Heinze G. A comparative investigation of methods for logistic regression with separated or nearly separated data. Stat Med 2006;25: 4216–26.