

Prognostic factors for perceived recovery or functional improvement in non-specific low back pain: secondary analyses of three randomized clinical trials

Pieter H. Helmhout · J. Bart Staal ·
Martijn W. Heymans · Chris C. Harts ·
Erik J. M. Hendriks · Rob A. de Bie

Received: 15 December 2008 / Revised: 30 September 2009 / Accepted: 10 December 2009 / Published online: 25 December 2009
© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract The objective of this study was to report on secondary analyses of a merged trial dataset aimed at exploring the potential importance of patient factors associated with clinically relevant improvements in non-acute, non-specific low back pain (LBP). From 273 predominantly male army workers (mean age 39 ± 10.5 years, range 20–56 years, 4 women) with LBP who were recruited in three randomized clinical trials, baseline individual patient factors, pain-related factors, work-related psychosocial factors, and psychological factors were evaluated as potential prognostic variables in a short-term (post-treatment) and a long-term logistic regression model (6 months after treatment). We found one dominant prognostic factor for improvement directly after treatment as well as 6 months later: baseline functional disability, expressed in Roland–Morris Disability Questionnaire scores. Baseline fear of movement, expressed in Tampa Scale for Kinesiophobia scores, had also significant prognostic value for

long-term improvement. Less strongly associated with the outcome, but also included in our final models, were supervisor social support and duration of complaints (short-term model), and co-worker social support and pain radiation (long-term model). Information about initial levels of functional disability and fear-avoidance behaviour can be of value in the treatment of patient populations with characteristics comparable to the current army study population (e.g., predominantly male, physically active, working, moderate but chronic back problems). Individuals at risk for poor long-term LBP recovery, i.e., individuals with high initial level of disability and prominent fear-avoidance behaviour, can be distinguished that may need additional cognitive-behavioural treatment.

Keywords Low back pain · Prognostic factors · Secondary analyses · Logistic regression · Multiple imputation

P. H. Helmhout (✉) · C. C. Harts
Department of Training Medicine and Training Physiology,
Personnel Command, Royal Netherlands Army, PO Box 90004,
3509 AA Utrecht, The Netherlands
e-mail: p.h.helmhout@mindef.nl

J. B. Staal · E. J. M. Hendriks · R. A. de Bie
Department of Epidemiology, School for Public Health
and Primary Care (Caphri), Maastricht University, Maastricht,
The Netherlands

M. W. Heymans
Institute for Research in Extramural Medicine,
VU University Medical Center, Amsterdam, The Netherlands

M. W. Heymans
Department of Methodology and Applied Biostatistics,
Institute of Health Sciences, VU University, Amsterdam,
The Netherlands

Introduction

Low back pain (LBP) is seen as a largely self-limiting health problem, with rapid improvement usually occurring within several weeks [43]. However, once the pain is showing itself as a recurrent and chronic phenomenon, it is mostly associated with long-term disability and, consequently, a significant socioeconomic burden: some 80% of the health care and social costs are attributed to the 10% of cases with chronic pain and disability [39].

Accurately, identifying individuals with a good or unfavourable prognosis amongst patients presenting with LBP is an important goal in current back pain research. Being able to predict prognosis of LBP patients based on pre-treatment assessment of patient characteristics may

lead to more realistic expectations of recovery as well as to more effective and efficient use of treatment modalities in the prevention of chronicity [32]. The identification of factors that can or cannot be modified in patients who are at risk for developing longstanding LBP may facilitate the selection of patients who will most likely benefit from targeted treatment. For example, if fear of movement appears to hinder a favourable prognosis in a certain subgroup of LBP patients, future exercises designed to mobilize or strengthen vulnerable lumbar body parts may become more effective in this target population when assisted by educational information and behavioural strategies regarding fear-related pitfalls [11].

Despite the fact that a considerable amount of research has been accumulated on a wide range of prognostic factors for LBP, inconsistencies amongst study results have limited the strength of conclusions. These inconsistencies have partly been attributed to methodological weaknesses of the studies involved, i.e., recruitment of heterogeneous cohorts in different settings and on different time-points; lack of an overarching conceptual framework; different use of measurements; model building with more variables than justified for the given number of observations; and/or incorrect use of statistical regression methods [1, 2, 8, 32, 43, 45]. Performing studies of sufficient statistical power on relatively homogeneous back-injured populations, preferably with a core set of measurements that are supported by the literature, may be of use for future reviews and meta-analyses that aim to identify those factors most strongly related to the onset and recurrence of LBP.

In recent years, we have conducted three randomized clinical trials to evaluate the effectiveness of different exercise modalities in an army working population with non-acute, non-specific LBP [24–27]. We provided isolated lumbar extensor strengthening versus mobilization versus general exercise therapy as our treatment interventions. Participants were 273 predominantly male soldiers from the Royal Netherlands Army with 4 weeks or more of low back complaints, who were referred to physiotherapy by the general practitioner of the military health centre. Consistent with prior evidence [5, 13, 37], neither of the exercise modalities seemed to offer incremental improvements over the other, up to 1 year post-treatment.

For several reasons, the data from these trials could be of value in secondary analyses aimed at identifying prognostic factors for clinically important LBP improvements. First, our study population can be considered relatively homogeneous. Inclusion and exclusion criteria were well-defined and comparable amongst trials. Subjects were selected from a patient group normally considered suitable for progressive resistance training, i.e., all patients with clinical contraindications (e.g., affected nerve root) were excluded. Participants were all professional military

employees and predominantly male (in total only four women were included), reflecting the vast male majority in our organization. Most participants were military recruits (younger population), military instructors (older population) or military staff personnel. All participants were working at the time of their inclusion, with, in many cases, similar physical and mental job demands. Overall, the large majority of our study population perceived their work as non-physical, despite the fact that working in uneasy or fixed positions and frequent lifting/carrying of heavy loads was reported in many cases. More than half of the participants were physically active in daily life despite their mostly moderate back complaints. This may be a reflection of the more-than-average physical attitude of military personnel in general; physical fitness is a critical aspect of military readiness and an inherent part of military service. Initial scores on a self-perceived health questionnaire (SF-36) that was used in two of our three trials indicated that our study population tend to attribute unspecific physical symptoms towards somatic disease more than towards mental issues [24, 25]. Furthermore, the study groups of our three trials were similar at baseline on most potential prognostic factors. The proportions of patients showing a favourable long-term outcome were also comparable amongst the three trials (71, 73, and 69%, respectively). Finally, the three trials had several common measures that could be used for pooled analyses. These measures have recently been recognized and included in a core set of factors for prospective cohorts in LBP [45].

The purpose of this paper is to report on secondary analyses of a merged trial dataset aimed at exploring the prognostic value of individual patient factors, pain-related factors, work-related psychosocial factors, and psychological factors, in non-acute, non-specific LBP. By identifying these prognostic factors, LBP management could be improved by adjusting current therapy concepts or by targeting therapies at those likely to gain the greatest benefit.

Materials and methods

Study design

We used a prospective cohort design by merging data from three randomized trials on the effectiveness of exercise therapy in individuals with non-specific LBP. Patients enrolled in the trials were randomized into an intervention group that received either an 8–12-week, high-intensive or low-intensive isolated lumbar extensor training program using specific training devices, or usual care that mainly consisted of general exercise therapy. Average numbers of treatment sessions varied from 8 to 14, depending on the program. Mean intervention times per treatment session

varied from 10 to 15 min (lumbar extensor training), and from 25 to 30 min (regular physiotherapy). An extensive description of the design and results of these trials have been published in recent years [24–27].

Subjects

The source population consisted of professional military employees of the Royal Netherlands Army ($n =$ approx. 23,000). In total, 273 predominantly male participants (mean age 39 ± 10.5 years, range 20–56 years, 4 women) were recruited during regular GP consulting hours as well as through advertisement in military media. None of the participants were performing combat-related activities during their follow-up in the trials. Inclusion criteria were: at least 4 weeks of continuous LBP or recurrent (at least three times a week) episodes of LBP; pain localized between posterior iliac crests and angulus inferior scapulae; availability during duty time; and willingness to abandon other treatment interventions for the lower back during the intervention period. Exclusion criteria were: received spinal surgery in the last 2 years; specific treatment for LBP in the last 4 weeks (e.g., physiotherapy, manual therapy); severe LBP which hindered performing maximal isometric strength efforts; specific LBP, defined as herniated disk, ankylosing spondylitis, spondylolisthesis, or relevant neurological diseases. All three trials used comparable methods to collect demographic and clinical information prior to randomization.

Prognostic factors

At baseline, directly after treatment and 6 months after treatment, respectively, patients completed a compound questionnaire containing, amongst others, the following items:

- functional disability, measured with the 24-item Roland–Morris Disability Questionnaire (RMDQ) [49];
- duration of LBP complaints at baseline, categorized as: 4–6 weeks; 6–12 weeks; 3–6 months; 6–12 months; 1 year or more. For analyses' purposes, we dichotomized this variable, using a cut-off point of 1 year to form balanced groups of patients with a shorter and longer duration of complaints, respectively;
- pain radiation or tingling in the legs at baseline (yes/no);
- fear of movement, measured with the validated Tampa Scale for Kinesiophobia (TSK) [34, 48, 62];
- psychological distress, using the 12-item General Health Questionnaire (GHQ) [20], and only measured in the third trial ($N = 127$);
- subscales 'Supervisor Social Support' and 'Co-worker Social Support' of the Job Content Questionnaire (JCQ) [31], only measured in the third trial ($N = 127$);

- degree of physical activity, measured in the first two trials with the criterion 'physically active for at least 30 min/day' [41], and, in the third trial, with the validated Short Questionnaire to Assess Health Enhancing Physical Activity (SQUASH) [65].

Baseline values of these eight variables (RMDQ, LBP duration, pain radiation, TSK, GHQ, JCQ subscales, physical activity) were included in the analyses as potential prognostic factors. The choice to include these factors in our analyses was based on the fact that (a) they were considered core prognostic factors in at least two out of three recent reviews on prospective cohorts in persistent LBP disability [8, 45, 52] or (b) they were considered relevant for the population under study, based on earlier experience (e.g., degree of physical activity).

Outcome

Patient improvement was selected as the dependent variable in our prognostic model. This variable was composed of changes in RMDQ scores and self-assessed changes in back complaints, post-treatment and after 6 months of follow-up, respectively. Self-assessed change in back complaints since the start of the treatment was scored on a percentual scale (0–100% improvement) in the first two trials [24, 25], or the seven-item Global Perceived Effect (GPE: 1 = completely recovered, 2 = much improved, 3 = slightly improved, 4 = no change, 5 = slightly worsened, 6 = much worsened, 7 = vastly worsened) in the third trial [3], respectively. The outcome variable was dichotomized into 'improved' and 'non-improved'. We defined 'improved' defined as subjects who met both the following criteria:

- 30% or more of improvement on the RMDQ;
- a score 'completely recovered' or 'much improved' on the GPE, or 20% or more of self-assessed improvement on the percentual scale.

These criteria were partly derived from recommendations by Jordan et al. [30] on clinically important differences in LBP, based on the RMDQ.

Statistical analyses

Model building

Prognostic variables and outcome variables with incomplete baseline and follow-up data were completed using the Multiple Imputation by Chained Equations (MICE) procedure [59]. In a multiple imputation procedure, each missing value is replaced by a set of multiple different values. These values are estimated using regression models

and all available data. We generated five multiply imputed datasets, according to Schafer [51].

The relationship between the outcome directly after treatment and 6 months after treatment, respectively, and each of the potential prognostic factors was individually calculated, using univariate logistic regression analyses. The allocated trial intervention was included in all analyses. Univariate odds ratios (ORs) were calculated to reflect the strength of each relationship, together with the corresponding 95% confidence interval (95% CI).

To determine which combination of factors was related to the outcome, we included all eight potential prognostic factors in a multivariable logistic regression model. This takes into account the ‘rule of thumb’ in logistic regression that the number of the less common of the two possible outcomes (in our case: ‘non-improved’ with, on average, 96 cases post-treatment and 89 cases at follow-up) divided by the maximum number of prognostic factors in the model should be at least 10 [42]. Backward regression analysis was applied to build each model, using a variable selection method that has recently been recommended by Wood and Royston [66]. With this method, backward regression is performed taking into account all imputed datasets. The initial regression model including all potential prognostic factors is fitted on each imputed dataset. Regression coefficients and related standard errors and p values are then estimated over all multiply imputed datasets using Rubin’s Rules [50].

Then, like with ‘normal’ backward regression, the variable with the highest p value is first omitted from the model. This smaller model is again fitted on each imputed dataset and, again, the variable with the largest p value is omitted. This procedure is repeated until all variables with a p value of <0.50 are retained in the model, following recommendations by Steyerberg et al. [55]. A liberal p value increases the changes of obtaining true predictors, limiting the bias in selected coefficients.

Model performance

The goodness-of-fit of each model was verified with the Hosmer–Lemeshow test [29]. A non-significant χ^2 value ($\alpha = 0.05$) in this test is indicative of a good model fit. In addition, we used residual regression diagnostics (Cook’s Distance, Leverage, Studentized Residual, and DFBeta) in revealing the effect on the estimated models of individual observations that are not adequately described by the model or that are highly influential on the model fit [15]. Moreover, we used collinearity diagnostics to check if factors were highly correlated [15]. A potential nonlinear behaviour of the continuous factors with the outcome was examined using restricted cubic spline functions and spline plots. Restricted cubic spline functions allow continuous

indicators to be fitted within the regression model without assuming a linear relation [23]. We did find a nonlinear relation for baseline RMDQ score and, therefore, included this variable in restricted cubic spline form in our model selection process. The spline variable was converted into a dummy variable (quartile categories with scores <4 , 4–7, 7–11, ≥ 11) in our final model to enhance clinical interpretation. We checked if the interpretation of the ORs would change if we fitted the models without the RMDQ spline function, which was not the case. All goodness-of-fit analyses were applied on the first imputed dataset, the results of which were comparable to those of the other datasets.

Two measures were used to further assess model performance: Nagelkerke’s R^2 and the C index. Nagelkerke’s R^2 (R_N^2) is an approximation of the explained variance (R^2) concept for the ordinary regression model. The C index, calculated as the area under the curve of the receiver operating characteristic (ROC) plot, represents the concordance between predicted probabilities and observed outcomes for all possible pairs of patients. It, therefore, indicates the discriminative ability of the logistic model. A C index of 1.0 indicates perfect discrimination while a value of 0.5 indicates that the model performs no better than chance alone. Prognostic models usually perform better in the patient sample that was used to build the model than in other new patient samples, due to optimism in regression coefficients and performance measures. To estimate the amount of optimism in the C index and explained variance, we used bootstrapping techniques [47]. The model performance indices were calculated on each of the five imputed dataset and then averaged.

Software

Imputation, backward selection, and bootstrapping were performed with R [47]. Diagnostic analyses were performed using SAS Version 9.1 (goodness-of-fit) and SPSS for Windows Version 15.0 (regression diagnostics, multicollinearity).

Results

The merged dataset of the three previous trials consisted of 273 subjects with non-acute, non-specific LBP. Table 1 shows the characteristics of these subjects at baseline. The percentage of missing data of the potential prognostic factors before imputing varied from 0 to 9% save those that were only measured in the third trial, i.e., psychological distress and supervisor/co-worker social support (53% missing overall). From the original dataset, 119 out of 225 subjects (53%) were labelled as ‘improved’ directly after treatment, and 131 out of 210 subjects (62%) at 6 months

Table 1 Baseline patient characteristics

Characteristics	Trial 1 (<i>n</i> = 81)	Trial 2 (<i>n</i> = 65)	Trial 3 (<i>n</i> = 127)	Total (<i>n</i> = 273)
Number of female participants	0	0	4	4
Age (years)	41 ± 9 [21–55]	43 ± 8 [25–58]	36 ± 11 [20–56]	39 ± 10 [20–58]
Physical job characteristics	(<i>n</i> = 81)	(<i>n</i> = 65)	(<i>n</i> = 127)	
Often working in fixed postures (%)	–	–	72	
Often working in uneasy postures (%)	–	–	41	
Often using repetitive movements in work (%)	–	–	32	
Often lifting or carrying heavy loads (>5 kg) in work (%)	–	–	41	
Perceiving work as physically heavy (%)	23	14	26	
Physically active (%)	64 (<i>n</i> = 70)	52 (<i>n</i> = 65)	58 (<i>n</i> = 127)	58 (<i>n</i> = 262)
Isometric back strength (N m)	218 ± 72 (<i>n</i> = 78)	223 ± 59 (<i>n</i> = 65)	214 ± 64 (<i>n</i> = 123)	217 ± 65 (<i>n</i> = 266)
Duration of complaints (%)	(<i>n</i> = 80)	(<i>n</i> = 59)	(<i>n</i> = 127)	(<i>n</i> = 248)
4–6 weeks	0	0	12	5
6–12 weeks	0	0	24	10
3–6 months	4	2	16	8
6–12 months	15	8	11	12
>12 months	81	90	37	65
Pain radiation (%)	23 (<i>n</i> = 79)	45 (<i>n</i> = 65)	26 (<i>n</i> = 114)	30 (<i>n</i> = 258)
Work absenteeism in last year because of LBP (%)	14 (<i>n</i> = 79)	34 (<i>n</i> = 65)	20 (<i>n</i> = 86)	22 (<i>n</i> = 230)
RMDQ score (0–24 points)	7.4 ± 4.8 [0–19]	6.3 ± 4.4 [0–18]	8.1 ± 4.6 [0–19]	7.5 ± 4.7 [0–19]
TSK score (17–68 points)	38.4 ± 6.7 [24–54]	36.5 ± 8.5 [22–52]	38.4 ± 6.6 [24–57]	38.0 ± 7.1 [22–57]
GHQ (0–12 point)	–	–	2.4 ± 2.5 [0–12]	2.4 ± 2.5 [0–12] (<i>n</i> = 127)
JCQ-S (4–16 points)	–	–	11.2 ± 1.9 [4–15]	11.2 ± 1.9 [4–15] (<i>n</i> = 127)
JCQ-C (4–16 points)	–	–	11.9 ± 1.3 [9–16]	11.9 ± 1.3 [9–16] (<i>n</i> = 127)

Values are mean ± SD [range] or as otherwise indicated; number of events per cell corresponds with total number per trial or as otherwise indicated

RMDQ Roland–Morris Disability Questionnaire, *GHQ* General Health Questionnaire, *JCQ-S* Job Content Questionnaire subscale ‘Supervisor Social Support’, *JCQ-C* Job Content Questionnaire subscale ‘Co-worker Social Support’

of follow-up. From the overall study population of 273 participants, only 3% (*N* = 9) had a zero score on the RMDQ at baseline. All nine subjects scored moderate to substantial self-assessed improvements in back complaints on either of the other scales that contribute to our outcome variable.

Table 2 shows the analyses of the individual prognostic factors, indicating that baseline RMDQ score between 8 and 11 was significantly associated with improvement in LBP, both directly (OR, 3.57) and 6 months after treatment (OR 4.22). Baseline RMDQ scores between 4 and 7 (OR 2.29) and scores of 11 and more (OR 2.53), respectively, as well as baseline TSK score (OR 0.97) were significantly associated with long-term improvement.

The multivariate analyses (see Table 3) showed a final post-treatment model that included four prognostic factors together explaining 12% of the variation in outcome: functional disability, fear of movement, supervisor social

support, and duration of complaints. All other factors were eliminated due to the *p* < 0.50 criterion for backward regression. The final long-term model consisted of the following five factors (16% explained variance): functional disability, fear of movement, psychological distress, co-worker social support, and pain radiation. In the post-treatment model, the prognostic factor most strongly associated with improvement was a baseline RMDQ score between 8 and 11 (OR 3.98). In the long-term model, baseline RMDQ scores (ORs 2.97–7.31) as well as baseline TSK score (OR 0.91) were the strongest prognostic factors.

Both models showed fairly good discriminative power with a *C* index of 0.68–0.70. Moreover, the global goodness-of-fit test was not rejected (see Table 3) and the regression diagnostics were within normal ranges indicating adequate model fit (not presented). Collinearity diagnostics showed that the assumption of ‘no multicollinearity’ was met in both models (not presented).

Table 2 Univariate analyses of baseline prognostic factors for improvement in LBP disability, post-treatment and at 6 months of follow-up, corrected for intervention

Variable (range) ^a	Post-treatment				Follow-up			
	<i>B</i> (SE)	<i>p</i> value	OR	95% CI	<i>B</i> (SE)	<i>p</i> value	OR	95% CI
Baseline RMDQ (0–24) ^b								
Scores 4–7	0.65 (0.42)	0.990	1.92	0.83–4.40	0.83 (0.38)	0.029	2.29	1.09–4.85
Scores 8–11	1.27 (0.43)	0.003	3.57	1.53–8.28	1.44 (0.47)	0.002	4.22	1.70–10.60
Score ≥11	0.47 (0.38)	0.216	1.60	0.76–3.34	0.93 (0.40)	0.023	2.53	1.14–5.61
Baseline TSK score (17–68)	–0.01 (0.02)	0.457	0.99	0.95–1.02	–0.04 (0.02)	0.035	1.02	0.92–0.99
Baseline GHQ (0–12)	0.03 (0.05)	0.535	1.03	0.93–1.15	0.06 (0.06)	0.366	1.06	0.93–1.20
Baseline JCQ-S score (4–16)	0.18 (0.10)	0.114	1.20	0.99–1.45	–0.06 (0.07)	0.434	1.07	0.82–1.09
Baseline JCQ-C score (4–16)	0.03 (0.12)	0.818	1.03	0.81–1.30	–0.13 (0.16)	0.459	0.88	0.64–1.21
Physically active	0.07 (0.29)	0.806	1.07	0.60–1.92	0.13 (0.32)	0.687	1.14	0.60–2.18
Duration of complaints	–0.50 (0.30)	0.103	0.61	0.33–1.12	–0.39 (0.31)	0.206	0.68	0.37–1.25
Pain radiation	–0.03 (0.31)	0.914	0.97	0.51–1.81	–0.12 (0.30)	0.889	0.92	0.71–1.61

B (SE) estimate with corresponding standard error, *OR* odds ratio, *95% CI* 95% confidence interval, *RMDQ* Roland–Morris Disability Questionnaire, *GHQ* General Health Questionnaire, *JCQ-S* Job Content Questionnaire subscale ‘Supervisor Social Support’, *JCQ-C* Job Content Questionnaire subscale ‘Co-worker Social Support’

^a For RMDQ, TSK, and GHQ, a negative estimate reflects improvement; for JCQ-S and JCQ-C, a positive estimate reflects improvement

^b Reference category: ‘RMDQ score < 4’

Table 3 Multivariate models of prognostic factors for improvement in LBP disability, post-treatment and at 6 months of follow-up, corrected for intervention

Variable (range) ^a	Post-treatment				Follow-up			
	<i>B</i> (SE)	<i>p</i> value	OR	95% CI	<i>B</i> (SE)	<i>p</i> value	OR	95% CI
Baseline RMDQ (0–24) ^b								
Scores 4–7	0.76 (0.44)	0.088	2.13	0.89–5.12	1.09 (0.43)	0.011	2.97	1.28–6.89
Scores 8–11	1.38 (0.47)	0.003	3.98	1.60–9.92	1.99 (0.53)	0.000	7.31	2.55–20.97
Score ≥11	0.66 (0.42)	0.114	1.93	0.85–4.35	1.65 (0.55)	0.004	5.18	1.72–15.56
Baseline TSK score (17–68)	–0.03 (0.02)	0.216	0.97	0.93–1.02	–0.09 (0.03)	0.000	0.91	0.87–0.96
Baseline GHQ (0–12)					0.08 (0.07)	0.227	1.09	0.95–1.24
Baseline JCQ-S score (4–16)	0.17 (0.11)	0.132	1.19	0.94–1.50				
Baseline JCQ-C score (4–16)					–0.17 (0.16)	0.298	0.84	0.60–1.19
Physically active								
Duration of complaints	–0.50 (0.33)	0.131	0.61	0.32–1.16				
Pain radiation					–0.24 (0.39)	0.544	0.79	0.36–1.72
<i>R</i> ² (%)	18		12 ^c		23		16 ^c	
<i>C</i> index	0.72		0.68 ^c		0.73		0.70 ^c	
Hosmer–Lemeshow	χ^2 value = 10.315 (<i>df</i> 8); <i>p</i> value = 0.244				χ^2 value = 8.837 (<i>df</i> 8); <i>p</i> value = 0.356			

B (SE) estimate with corresponding standard error, *OR* odds ratio, *95% CI* 95% confidence interval, *RMDQ* Roland–Morris Disability Questionnaire, *GHQ* General Health Questionnaire, *JCQ-S* Job Content Questionnaire subscale ‘Supervisor Social Support’, *JCQ-C* Job Content Questionnaire subscale ‘Co-worker Social Support’, *R*² Nagelkerke’s *R*²

^a For RMDQ, TSK, and GHQ, a negative estimate reflects improvement; for JCQ-S and JCQ-C, a positive estimate reflects improvement

^b Reference category: ‘RMDQ score < 4’

^c Bootstrap-corrected values

Discussion

The aim of this study was to assess the relative importance of individual patient factors, pain-related factors, work-related psychosocial factors, and psychological factors, in explaining self-reported clinically important improvements in LBP complaints. Overall, we found one dominant prognostic factor for improvement directly after treatment as well as 6 months later: functional disability, more specifically intermediate RMDQ score at baseline. Fear of movement had also significant prognostic value for long-term improvement. These factors were also found to be individually related to the outcome in the univariate analyses. Less strongly associated with the outcome, but also included in our final multivariate models, were supervisor social support and duration of complaints (short-term model), and co-worker social support and pain radiation (long-term model).

The consistent appearance of the two strongest prognostic factors, functional disability and fear of movement, in our final models is in line with earlier prognostic LBP studies, in the sense that they are considered as important and independent determinants of many different LBP outcomes (e.g., remitting pain, sub-acute or chronic disability, failed or delayed recovery from short-term LBP, long-term compensation status, time to return-to-work) in various study settings (e.g., primary care, specialist back clinic, occupational health care, mailing survey). Concordant with other studies on LBP [6, 7, 12, 16, 40], functional disability (baseline RMDQ) was highly predictable in our models. Table 3 illustrates that midrange scores on the RMDQ (score 8–11) and, to a lesser extent, high range values (≥ 11) had stronger prognostic value than low range values (4–7). Possibly, this finding can be explained by a ‘law of diminishing returns’ phenomenon, i.e., individuals with high levels of functional disability at baseline have a higher potential to improve than those with low disability levels. Nordin et al. [40] have reported that LBP patients with severe functional disability according to the Oswestry Disability Index were found to return-to-work later, possibly due to greater episode severity and/or a higher perception of being ‘sick’. The fact that functional disability was our strongest prognostic factor was not a surprising finding, since changes in RMDQ values were part of the compound variable (improved vs. not improved) we constructed as our outcome measure. But also in other studies, using different outcomes (e.g., work compensation status, time to return-to-work), baseline functional disability appeared to have prognostic value [16, 61]. We have chosen an outcome measure that was partly based on the RMDQ, since we consider its dimension ‘functional disability due to perceived back pain’ as more relevant for our mainly chronically injured, working population than indicators of the

symptom itself, i.e., the degree of back pain. The treatments in our trials were predominantly aimed at restoring (physical) functioning, not at reducing the pain.

Our results support the conclusion of others that fear of movement has prognostic value in long-term disability [17, 19, 44, 56]. A positive influence of exercise on fear of movement behaviour has been reported in the literature [4, 11]. Exercises with fear-avoidance-based principles are thought to give additional benefits for subgroups of LBP patients with high levels of fear-avoidance beliefs [33, 58]. All three intervention modalities that were used in our trials consisted of exercise training, but none of them included specific fear-avoidance strategies. Our analyses suggest that fear-avoidance issues could have additional value when designing exercise treatment modalities in a LBP population that is already relatively physically active despite the presence of back pain.

Other factors in our models, which were less strongly associated with the outcome, were supervisor social support (short term), duration of complaints (short term), psychological distress (long term), co-worker social support (long term), and pain radiation (long term). The fact that supervisor social support was associated to short-term LBP improvement may very well be a reflection of the strong decisive influence of superiors in our hierarchical military organization on the participation of soldiers in duty time in activities such as our research program. We used different definitions of duration of complaints in our trials (current complaints vs. time since first LBP episode), all reflecting a different dimension of the history of the LBP symptoms. From the literature, we know that a previous history of LBP is highly predictive of persistent symptoms, suggesting that people with one or more previous episodes are likely to have future multiple episodes [64]. Psychological distress was another long-term factor in our model. Psychological distress has been recognized as a prognostic factor for both the onset and recurrence of LBP in many different populations, ranging from healthy young or middle-aged adults [14, 36] to patients in primary care settings [22, 35, 57] and to individuals with work-related back injuries [9, 16]. In line with our study findings, Gheldof et al. [19] found that pain radiation was a significant risk factor for the development of long-term LBP in a population of predominantly male industrial workers. This risk was reduced by social support of co-workers.

Physical activity was the only factor that was removed from our regression models. There is inconsistent evidence in the literature about whether physically active leisure time versus sedentary activity has influence on the development of musculoskeletal morbidity in general and LBP in particular [28]. Some recent studies show associations between chronic LBP and physical activity or aerobic fitness level, respectively [38, 53], but others do not [46].

In general, we were able to generate well-fitted but rather weak-performing prognostic models, with percentages of explained variance (12–16%) that are comparable or somewhat less favourable to those of other prognostic models in the same field of research [2, 10, 18, 19, 54, 58, 60]. This may partly be explained by the fact that other than the chosen prognostic indicators, such as pain severity or job satisfaction, may play an important role in our population. However, in most cohort studies on non-specific LBP, baseline factors only account for moderate amounts of variance in LBP outcomes, typically around 30% [45]. The large unexplained variances are, thus, most likely a reflection of the fact that an individual's course of non-specific LBP is highly complex, affected by interacting factors that probably cover the whole spectrum of the biopsychosocial model of pain and disability and that present themselves in different time phases of the process [63].

There are some limitations in our study that must be recognized. First, our findings need to be interpreted with some caution, because the analyses were based on data from randomized clinical trials. These trials were initially not designed to identify prognostic factors for LBP improvement. The merged dataset may, therefore, still have insufficient power to detect these prognostic factors, despite our efforts to prevent overfitting problems by matching the number of potential prognostic factors to the number of outcome events. However, the confidence intervals of all but the dummy variables (RMDQ, intervention) were small, indicating that our study findings are robust. Second, due to the fact that one of the trials had a relatively short follow-up period, we were not able to measure effects at longer than 6 months after treatment. Moreover, in neither of the studies we used a true control group that could have reflected the natural course of LBP. In other words, all participants had experienced exercise therapy through our research program in the previous year and should be regarded as such when extrapolating the results of this study to other populations. This brings us to the external validity of our results. The population under study (male soldiers) is clearly not a representation of those presenting most commonly in primary care. Obviously, one should be careful in generalizing the results of this study to other samples, such as white-collar or female workers.

The population under study can be seen as a rather homogeneous group of individuals, consisting of young and middle-aged male subjects in their active working period (87% between 25 and 55 years of age) who are not overly work-disabled by their condition despite incidentally reported high disability scores. In this respect, potential confounding influences of individual patient factors and work-related factors that have not been addressed in our analyses are expected to be low, which can be seen as strength of our study. Moreover, we included a number of

variables that have been identified as potentially important prognostic factors for LBP in several relevant areas (individual, work-related psychosocial, psychological) and that were measured with standardized, validated instruments. Finally, using state-of-the-art statistical techniques for model building and evaluation of model performance, we have tried to prevent substantial shortcomings that have frequently been addressed in the use and report of logistic regression analyses in medical research, such as overfitting, assumption checks, and unrestricted use of automated variable selection [7].

Not all trials collected the same prognostic factors. Consequently, after pooling the data, the percentage of missing data in three variables was around 50%. For these situations, multiple imputation produces valid results under the missing at random assumption, i.e., missings can be explained by the available data in the dataset since they are not dependent on the values of the variables themselves. This latter statement holds in our study because there was no specific reason in each of the trials to exclude variables. Furthermore, most variables in our study have shown to be important in LBP prognosis, which means that all variables could be used in the imputation model to estimate the missing values. Several studies have shown that with multiple imputation valid results can be obtained when around 50% of the data are missing [21].

Conclusions

This study implies that it can be of use for clinical practitioners to gather pre-treatment patient information, in particular on individual levels of functional disability and fear-avoidance behaviour, in patient populations with characteristics comparable to ours (e.g., predominantly male, physically active, working, moderate but chronic back problems). By doing so, individuals at risk for poor long-term LBP recovery, i.e., individuals with high initial level of disability and prominent fear-avoidance behaviour, can be distinguished that may need additional cognitive-behavioural treatment. Further research is warranted to find out if this strategy actually leads to a higher improvement rate.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

1. Bagley SC, White H, Golomb BA (2001) Logistic regression in the medical literature: standards for use and reporting, with particular attention to one medical domain. *J Clin Epidemiol* 54(10):979–985

2. Bekkering GE, Hendriks HJ, van Tulder MW et al (2005) Prognostic factors for low back pain in patients referred for physiotherapy: comparing outcomes and varying modeling techniques. *Spine* 30(16):1881–1886
3. Beurskens AJ, de Vet HC, Koke AJ (1996) Responsiveness of functional status in low back pain: a comparison of different instruments. *Pain* 65(1):71–76
4. Brox JI, Sorensen R, Friis A et al (2003) Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. *Spine* 28(17):1913–1921
5. Cairns MC, Foster NE, Wright C (2006) Randomized controlled trial of specific spinal stabilization exercises and conventional physiotherapy for recurrent low back pain. *Spine* 31(19):E670–E681
6. 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(1):115–120
7. Coste J, Lefrancois G, Guillemin F et al (2004) Prognosis and quality of life in patients with acute low back pain: insights from a comprehensive inception cohort study. *Arthritis Rheum* 51(2):168–176
8. Crook J, Milner R, Schultz IZ et al (2002) Determinants of occupational disability following a low back injury: a critical review of the literature. *J Occup Rehabil* 12(4):277–295
9. Crook J, Moldofsky H, Shannon H (1998) Determinants of disability after a work related musculoskeletal injury. *J Rheumatol* 25(8):1570–1577
10. Dionne CE, Koepsell TD, Von Korff M et al (1997) Predicting long-term functional limitations among back pain patients in primary care settings. *J Clin Epidemiol* 50(1):31–43
11. Elfving B, Andersson T, Grooten WJ (2007) Low levels of physical activity in back pain patients are associated with high levels of fear-avoidance beliefs and pain catastrophizing. *Physiother Res Int* 12(1):14–24
12. Enthoven P, Skargren E, Carstensen J et al (2006) Predictive factors for 1-year and 5-year outcome for disability in a working population of patients with low back pain treated in primary care. *Pain* 122(1–2):137–144
13. Ferreira ML, Ferreira PH, Latimer J et al (2007) Comparison of general exercise, motor control exercise and spinal manipulative therapy for chronic low back pain: a randomized trial. *Pain* 131(1–2):31–37
14. Feyer AM, Herbison P, Williamson AM et al (2000) The role of physical and psychological factors in occupational low back pain: a prospective cohort study. *Occup Environ Med* 57(2):116–120
15. Field A (2005) *Discovering statistics using SPSS*. SAGE Publications Ltd, London
16. Fransen M, Woodward M, Norton R et al (2002) Risk factors associated with the transition from acute to chronic occupational back pain. *Spine* 27(1):92–98
17. Fritz JM, George SZ (2002) Identifying psychosocial variables in patients with acute work-related low back pain: the importance of fear-avoidance beliefs. *Phys Ther* 82(10):973–983
18. George SZ, Fritz JM, Childs JD et al (2006) Sex differences in predictors of outcome in selected physical therapy interventions for acute low back pain. *J Orthop Sports Phys Ther* 36(6):354–363
19. Gheldof EL, Vinck J, Vlaeyen JW et al (2007) Development of and recovery from short- and long-term low back pain in occupational settings: a prospective cohort study. *Eur J Pain* 11(8):841–854
20. Goldberg DP, Williams P (1988) *A user's guide to the general health questionnaire*. NFER-Nelson, Windsor
21. Greenland S, Finkle WD (1995) A critical look at methods for handling missing covariates in epidemiologic regression analyses. *Am J Epidemiol* 142(12):1255–1264
22. Grotle M, Brox JI, Glomsrod B et al (2007) Prognostic factors in first-time care seekers due to acute low back pain. *Eur J Pain* 11(3):290–298
23. Harrell FE Jr (2001) *Regression modeling strategies*. Springer, Berlin
24. Harts CC, Helmhout PH, de Bie RA et al (2008) A high-intensity lumbar extensor strengthening program is little better than a low-intensity program or a waiting list control group for chronic low back pain: a randomised clinical trial. *Aust J Physiother* 54(1):23–31
25. Helmhout PH, Harts CC, Staal JB et al (2004) Comparison of a high-intensity and a low-intensity lumbar extensor training program as minimal intervention treatment in low back pain: a randomized trial. *Eur Spine J* 13(6):537–547
26. Helmhout PH, Harts CC, Staal JB et al (2004) Rationale and design of a multicenter randomized controlled trial on a “minimal intervention” in Dutch army personnel with nonspecific low back pain [ISRCTN19334317]. *BMC Musculoskelet Disord* 5(1):40
27. Helmhout PH, Harts CC, Viechtbauer W et al (2008) Isolated lumbar extensor strengthening versus regular physical therapy in an army working population with nonacute low back pain: a randomized controlled trial. *Arch Phys Med Rehabil* 2008 July 31 [Epub ahead of print]
28. Hildebrandt VH, Bongers PM, Dul J et al (2000) The relationship between leisure time, physical activities and musculoskeletal symptoms and disability in worker populations. *Int Arch Occup Environ Health* 73(8):507–518
29. Hosmer D, Lemeshow S (1980) A goodness-of-fit test for the multiple logistic regression model. *Commun Stat A10*:1043–1069
30. Jordan K, Dunn KM, Lewis M et al (2006) A minimal clinically important difference was derived for the Roland–Morris Disability Questionnaire for low back pain. *J Clin Epidemiol* 59(1):45–52
31. Karasek R, Brisson C, Kawakami N et al (1998) The Job Content Questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. *J Occup Health Psychol* 3(4):322–355
32. Kent PM, Keating JL (2008) Can we predict poor recovery from recent-onset nonspecific low back pain? A systematic review. *Manual Ther* 13(1):12–28
33. Klaber Moffett JA, Carr J, Howarth E (2004) High fear-avoiders of physical activity benefit from an exercise program for patients with back pain. *Spine* 29(11):1167–1172 discussion 1173
34. Kori SH, Miller RP, Dodd DD (1990) Kinesiophobia: a new view of chronic back pain behavior. *Pain Manag Jan/Feb*:35–43
35. Langworthy JM, Breen AC (2007) Psychosocial factors and their predictive value in chiropractic patients with low back pain: a prospective inception cohort study. *Chiropr Osteopat* 15:5
36. Linton SJ (2005) Do psychological factors increase the risk for back pain in the general population in both a cross-sectional and prospective analysis? *Eur J Pain* 9(4):355–361
37. Mannion AF, Muntener M, Taimela S et al (1999) A randomized clinical trial of three active therapies for chronic low back pain. *Spine* 24(23):2435–2448
38. Morken T, Mageroy N, Moen BE (2007) Physical activity is associated with a low prevalence of musculoskeletal disorders in the Royal Norwegian Navy: a cross sectional study. *BMC Musculoskelet Disord* 8:56
39. Nachemson AL, Jonsson E (2000) *Neck and back pain: the scientific evidence of causes, diagnosis and treatment*. Lippincott Williams & Wilkins, Philadelphia
40. Nordin M, Skovron M, Hiebert R et al (1997) Early predictors of delayed return to work in patients with low back pain. *J Musculoskelet Pain* 5(2):5–27
41. Pate RR, Pratt M, Blair SN et al (1995) Physical activity and public health. A recommendation from the Center for Disease

- Control and Prevention and the American College of Sports Medicine. *JAMA* 273:402–407
42. Peduzzi P, Concato J, Kemper E et al (1996) A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 49(12):1373–1379
 43. Pengel LH, Herbert RD, Maher CG et al (2003) Acute low back pain: systematic review of its prognosis. *BMJ* 327(7410):323
 44. Picavet HS, Vlaeyen JW, Schouten JS (2002) Pain catastrophizing and kinesiophobia: predictors of chronic low back pain. *Am J Epidemiol* 156(11):1028–1034
 45. Pincus T, Santos R, Breen A et al (2008) A review and proposal for a core set of factors for prospective cohorts in low back pain: a consensus statement. *Arthritis Rheum* 59(1):14–24
 46. Rasmussen-Barr E, Lundqvist L, Nilsson-Wikmar L et al (2008) Aerobic fitness in patients at work despite recurrent low back pain: a cross-sectional study with healthy age- and gender-matched controls. *J Rehabil Med* 40(5):359–365
 47. R Development Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available at: <http://www.R-project.org>. Accessed 18 July 2008
 48. Roelofs J, Goubert L, Peters ML et al (2004) The Tampa Scale for Kinesiophobia: further examination of psychometric properties in patients with chronic low back pain and fibromyalgia. *Eur J Pain* 8(5):495–502
 49. Roland M, Morris R (1983) A study of the natural history of back pain. Part I. Development of a reliable and sensitive measure of disability in low-back pain. *Spine* 8(2):141–144
 50. Rubin DB (1987) Multiple imputation for nonresponse in surveys. Wiley, New York
 51. Schafer JL (1999) Multiple imputation: a primer. *Stat Methods Res* 8(1):3–15
 52. Shaw WS, Pransky G, Fitzgerald TE (2001) Early prognosis for low back disability: intervention strategies for health care providers. *Disabil Rehabil* 23(18):815–828
 53. Smeets RJ, Wittink H, Hidding A et al (2006) Do patients with chronic low back pain have a lower level of aerobic fitness than healthy controls? Are pain, disability, fear of injury, working status, or level of leisure time activity associated with the difference in aerobic fitness level? *Spine* 31(1):90–97 discussion 98
 54. Steenstra IA, Koopman FS, Knol DL et al (2005) Prognostic factors for duration of sick leave due to low-back pain in Dutch health care professionals. *J Occup Rehabil* 15(4):591–605
 55. Steyerberg EW, Eijkemans MJ, Harrell FE Jr et al (2001) Prognostic modeling with logistic regression analysis: in search of a sensible strategy in small data sets. *Med Decis Making* 21(1):45–56
 56. Swinkels-Meewisse IE, Roelofs J, Verbeek AL et al (2003) Fear of movement/(re)injury, disability and participation in acute low back pain. *Pain* 105(1–2):371–379
 57. Thomas E, Silman AJ, Croft PR et al (1999) Predicting who develops chronic low back pain in primary care: a prospective study. *BMJ* 318(7199):1662–1667
 58. Underwood MR, Morton V, Farrin A (2007) Do baseline characteristics predict response to treatment for low back pain? Secondary analysis of the UK BEAM dataset [ISRCTN32683578]. *Rheumatology (Oxford)* 46(8):1297–1302
 59. Van Buuren S, Oudshoorn K (1999) Flexible multivariate imputation by MICE. Technical report TNO quality of life. Leiden, The Netherlands
 60. Van der Hulst M, Vollenbroek-Hutten MM, Groothuis-Oudshoorn KG et al (2008) Multidisciplinary rehabilitation treatment of patients with chronic low back pain: a prognostic model for its outcome. *Clin J Pain* 24(5):421–430
 61. Van der Weide WE, Verbeek JH, Salle HJ et al (1999) Prognostic factors for chronic disability from acute low-back pain in occupational health care. *Scand J Work Environ Health* 25(1):50–56
 62. Vlaeyen JW, Kole-Snijders AM, Boeren RG, van Eek H (1995) Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. *Pain* 62:363–372
 63. Waddell G (1998) The back pain revolution. Churchill Livingstone, London
 43. Waddell G, Burton AK (2001) Occupational health guidelines for the management of low back pain at work: evidence review. *Occup Med (Lond)* 51(2):124–135
 65. Wendel-Vos GC, Schuit AJ, Saris WH et al (2003) Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. *J Clin Epidemiol* 56(12):1163–1169
 66. Wood AM, White IR, Royston P (2008) How should variable selection be performed with multiply imputed data? *Stat Med* 27(17):3227–3246