Manual Therapy 20 (2015) 96–102

Contents lists available at ScienceDirect

Manual Therapy

journal homepage: www.elsevier.com/math

Original article

Prognostic factors and course for successful clinical outcome quality of life and patients' perceived effect after a cognitive behavior therapy for chronic non-specific low back pain: A 12-months prospective study

Karin Verkerk [a, b, c, *], Pim A.J. Luijsterburg [c], Martijn W. Heymans [d, e, f], Inge Ronchetti [b], Harald S. Miedema [a], Bart W. Koes [c], Annelies Pool-Goudzwaard [g]

[a] Rotterdam University of Applied Sciences, Rochussenstraat 198, 3015 EK Rotterdam, The Netherlands
[b] Spine & Joint Centre, Rotterdam, The Netherlands
[c] Department of General Practice, Erasmus MC, University Medical Center, Rotterdam, The Netherlands
[d] EMGO Institute for Health and Care Research, VU University Medical Center, Amsterdam, The Netherlands
[e] Department of Methodology and Applied Biostatistics, VU University, Institute for Health Sciences, Amsterdam, The Netherlands
[f] Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands
[g] Department of Neuroscience, Erasmus MC, University Medical Center, Rotterdam, The Netherlands

ARTICLE INFO

Article history:
Received 4 October 2013
Received in revised form
5 July 2014
Accepted 7 July 2014

Keywords:
Chronic non-specific low back pain
Course
Prognosis
Psychological factors

ABSTRACT

This study investigates the clinical course of and prognostic factors for quality of life (Short Form 36 items Health survey (SF-36)) and global perceived effect (GPE) in patients treated for chronic non-specific low back pain at 5 and 12-months follow-up. Data from a prospective cohort (n = 1760) of a rehabilitation center were used, where patients followed a 2-months cognitive behavior treatment. The outcome 'improvement in quality of life (SF-36)' was defined as a 10% increase in score on the SF-36 at follow-up compared with baseline. On the GPE scale, patients who indicated to be 'much improved' were coded as 'clinically improved'. Multivariable logistic regression analysis included 23 baseline characteristics. At 5-months follow-up, scores on the SF-36 Mental Component Scale (SF-36; MCS) and the Physical Component Scale (SF-36; PCS) had increased from 46.6 (SD 10.3) to 50.4 (SD 9.8) and from 31.9 (SD 7.1) to 46.6 (SD 10.3), respectively. At 5-months follow-up, 53.0% of the patients reported clinical improvement (GPE) which increased to 60.3% at 12-months follow-up. The 10% improvement in quality of life (SF-36 MCS) at 5-months follow-up was associated with patient characteristics and psychological factors. At 5-months follow-up, the 10% improvement in quality of life (SF-36 PCS) and GPE was associated with patient characteristics, physical examination, work-related factors and psychological factors; for GPE, an association was also found with clinical status. At 12-months follow-up GPE was associated with patient characteristics, clinical status, physical examination and work-related factors. The next phase in this prognostic research is external validation of these results.

1. Introduction

Chronic non-specific low back pain (CNSLBP) is one of the most prevalent health problems (Heneweer et al., 2007). Although it is known that physical, psychosocial and personal factors play a role, the way they interact with each other remains unclear. Several prognostic models for non-specific low back pain have been described; however, the prognostic factors varied depending on the choice of, for example, the prognostic variables, outcome definition, or the stage of pain (e.g. acute, sub-acute or chronic) (Kent and Keating, 2008; Costa Lda et al., 2009; Verkerk et al., 2012). A recent systematic review focusing on musculoskeletal complaints considered relevant for physical therapists in primary care, reported that the available prediction models are not yet ready to be applied in clinical practice because of their preliminary stage of development (van Oort et al., 2012). Also, the available models for

* Corresponding author. Rotterdam University of Applied Sciences, Health Care, Rochussenstraat 198, 3015 EK Rotterdam, The Netherlands. Tel.: +31 10 7945153, +31 102260187; fax: +31 10 7944023.
E-mail addresses: k.verkerk@hr.nl, karinsomt@gmail.com (K. Verkerk).

http://dx.doi.org/10.1016/j.math.2014.07.003
1356-689X/

back pain patients need external validation and impact evaluation before applying them in daily practice (van Oort et al., 2012). Compared to patients with (sub) acute NSLBP, patients with CNSLBP are the least investigated regarding their course and prognosis, especially in relation to the outcomes 'quality of life' and 'global perceived effect' (GPE) (Verkerk et al., 2012). Therefore, clinicians and researchers increasingly recognize the importance of such patient-reported outcome measures in the evaluation of the effectiveness of treatment, prognosis or course of CNSLBP (Bombardier, 2000).

Achieving and maintaining the best possible quality of life is a primary goal of care and several questionnaires are available to measure this item, including the Short Form 36-items Health Survey (SF-36) (Aktekin et al., 2009). With regard to evaluating GPE, the patient can be asked to rate how much their condition (i.e. important aspects of recovery) has improved or deteriorated since some predefined time point (Kamper et al., 2010). The present study was designed to investigate the course of and identify prognostic factors (with internal validation) for quality of life and GPE in patients treated for CNSLBP.

2. Methods

2.1. Population

Patients were recruited between January 2003 and December 2008 in a prospective cohort study from a multidisciplinary outpatient rehabilitation clinic the Spine & Joint Centre (SJC; Rotterdam, The Netherlands). The Medical Ethics Committee of SJC approved the study protocol and all patients provided informed consent. Details on the study design are described elsewhere (Verkerk et al., 2011). Inclusion criteria were: 1) men and women aged $\geq$18 years; 2) having CNSLBP defined as a duration of LBP for $\geq$3 months; 3) having persistent low back complaints despite of treatment in primary and/or secondary care.

Exclusion criteria were insufficient knowledge of the Dutch language; signs indicating radiculopathy, asymmetric Achilles tendon reflex and/or (passive) straight leg raise test restricted by pain in the lower leg; positive MRI findings for disc herniation; recent (<6 months) fracture, neoplasm or recent previous surgery (<6 months) of the lumbar spine, the pelvic girdle, the hip joint, or the femur; specific causes such as ankylosing spondylitis and systemic disease of the locomotor system; and being pregnant or $\leq$6 months post-partum at the moment of consultation.

A total of 2545 patients [mean age 40.4 (10.9) years; 73.3% women] visited the SJC for an intake consultation between 2003 and 2008, but 785 patients [mean age 41.3 (11.5) years; 70.3% women] decided not to start therapy (e.g. only wanted consultation, diagnose, advise, referred to another specialist, decided later not to come). Data were collected at baseline (n = 1760) and at 2 (n = 1696), 5 (n = 1564) and 12 (n = 965) months-follow-up (Verkerk et al., 2011) during regular daily care at the SJC.

2.2. Measurements

2.2.1. Outcome measures and defining recovery

To determine the course of quality of life in patients with CNSLBP the SF-36 was used and, at 5 months, represented by the two SF-36 domains the Mental Component Scale (SF-36; MCS) and the Physical Component Scale (SF-36; PCS), both ranging from 0 to 100 (high quality of life) (Gandek et al., 1998; Walsh et al., 2003; Davidson et al., 2004; Gandek et al., 2004). Clinical improvement was measured at 2, 5 and 12-months follow-up with the GPE score, which consists of a 5-point scale on global change (1 = much improved, 2 = slightly improved, 3 = no change, 4 = slightly worsened, 5 = much worsened) (Ostelo and de Vet, 2005). The two instruments have shown to be reliable and valid (Walsh et al., 2003; Hagg et al., 2003b; Davidson et al., 2004; Gandek et al., 2004; Kamper et al., 2010).

Recovery was defined as a 10% improvement on the MCS or PCS compared to baseline. The scale was dichotomized into 'no improvement in MCS or PCS' and 'improvement in MCS or PCS' based on an increase of 10% at follow-up compared to the baseline value; we considered this to be a clinically relevant difference. A clinically relevant improvement for these scales has not yet been defined, but beside empirical evidence an expert clinical interpretation and judgment is of value. By expert opinion the most appropriate value for questionnaires on 'quality of life' is 10% since the changes are smaller than the more common outcomes measures on pain and disability. The SF-36 was only followed up to 5 months because this was done electronically at the SJC. The predefined time point for the GPE score (Kamper et al., 2010) was measured following 2 months of therapy at the SJC. In addition, patients judged their own improvement compared with this previous measurement, at 5 and 12-months follow-up. Patients who indicated 'much improved' were coded 'clinically improved' and patients who indicated 'slightly improved', 'no change', 'slightly worsened' or 'much worsened' were coded as 'clinically not improved' (Ostelo and de Vet, 2005).

2.2.2. Potential prognostic factors

The selection of relevant prognostic factors was performed in two steps: 1) the literature on prognosis for CNSLBP and quality of life and GPE were reviewed (Verkerk et al., 2012), and 2) a clinical group of 8 experts on CNSLBP composed a list of 23 of the 47 potential prognostic factors. All factors were retrieved from step 1 (with exception of the factor previous rehabilitation) in combination of the available variables at the SJC. Using the Policy Delphi method (scored on a 4-point Likert scale ranging from 1 = very important to 4 = not important) (Verhagen et al., 1998; Snyder-Halpern, 2001), there were 3 rounds and each time the responses were aggregated, tabulated, summarized, and returned to the experts. In the third round the experts were asked to decide whether to keep or remove the factor from the list, through consensus meeting. The final list consisted of factors that were included by at least 80% consensus. Using these 23 variables, in the analysis we complied with the rule of at least 10 events per variable (which avoids incorrect estimation of variables), we had to restrict the total number of potential prognostic factors (Peduzzi et al., 1996) (Box 1). We described the baseline values of these 23 potential prognostic factors in Table 1 in several domains (e.g. patients characteristics) to be transparent with other studies (Bombardier, 2000; Pincus et al., 2008; Kamper et al., 2010; Verkerk et al., 2012) studying on outcome measurements and clinical improvement. The excluded prognostic factors can be obtained from the first author.

2.3. Treatment at the Spine & Joint Centre

The multidisciplinary treatment at the SJC centre used a biopsychosocial approach consisting of 16 sessions of 3 h each during a 2-month period (total of 48 h). Patients were coached by a multidisciplinary team (e.g. a physical therapist, physician, health scientist, psychologist) (Verkerk et al., 2011).

2.4. Data analysis

2.4.1. Course of quality of life and GPE

Descriptive analysis was performed to describe the course of quality of life (SF-36; PCS and MCS) and GPE in CNSLBP patients according to their characteristics.

Box 1
. The 23 potential prognostic factors.

Continuous variables
1 Age (years)
2 Duration of back pain in years
3 Present pain intensity (VAS: 0–100 mm)
4 Degree of present fatigue (VAS: 0–100 mm)
5 Quebec Back Pain Disability scale (QBPDS: 0–100)
6 Tampa scale for kinesiophobia (TSK, 17–68)
7 Short-form health survey 36 (SF-36); Physical Component Scale (PCS) (range 0 "low quality of life" – 100 points)
8 Short-form health survey 36 (SF-36); Mental Component Scale (MCS) (range 0 "low quality of life" – 100 points)
9 Symptom Checklist 90 (SCL-90; item 9; psychoneuroticism)
10 Work participation (0–100%)
11 B200 Isostation (strength back extension in Newton)

Categorical variables
12 Body Mass Index (BMI ≤24.9/25–29.9/≥30 kg/m^2)
13 Cause of back pain (accident movement; after physical load; during pregnancy or after delivery; unknown; surgery pelvis/back or HNP)
14 Course of pain in the previous 3 months (stable; increased; decreased)
15 Duration of walking (0–15/16–30/31–60/>61 min)
16 Duration of sitting (0–15/16–30/31–60/>61 min)
17 Duration of standing (0–15/16–30/31–60/>61 min)

Dichotomized variables
18 Gender (female/male)
19 Co-morbidity (no versus having one or more co-morbidities)
20 Marital status (being alone versus being married/living with one adult)
21 Level of education (<high school versus ≥ high school /university)
22 Previous rehabilitation treatment (no versus one or more previous rehabilitation treatments)
23 Sickness benefit (no versus all kinds of benefits from the government or employer)

The percentage of patients defined as recovered based on a 10% improvement of the MCS and PCS at 2 and 5-months follow-up compared to baseline, was calculated. This was also done for GPE, 'clinically improved' versus 'not clinically improved', at 2, 5 and 12-months follow-up.

2.4.2. Model development

First, eligible prognostic factors were identified which were highly correlated ($r > 0.8$). This was the case for the B200 Isostation (strength in flexion, extension, lateroflexion, rotation) and the SCL-90 (items 1–8). Only the B200 extension and the total score item 9 of the SCL-90 were included in the analysis (van Buuren, 2012). The continuous factors were checked for linearity using spline regression curves which revealed a non-linear relationship between body mass index (BMI) and the PCS, MCS or GPE. Therefore, BMI was changed into a categorical variable, and also used for the present study and the presented outcomes.

To develop our prognostic model, multivariable logistic regression analysis was performed (Harrell, 2001; Royston et al., 2009; Moons et al., 2009a, 2009b). Regarding missing values, we applied multiple imputation of 5 datasets (van Buuren, 2012). Regression equations are used to estimate the missing values. Results of 5 imputed datasets were compared when 40 imputed datasets are used to see if the results would change; this number of 40 was used because in the initial model selection 45.2% of the patients at 12 months ($n = 795$) was missing (loss-to-follow up). Because the results were similar, 5 imputed datasets were used as the primary analysis methods. We also compared the results with complete case analysis (CCA), i.e. all patients with missing data were excluded from the analyses (Steyerberg et al., 2004; van Buuren, 2012).

To develop our prognostic model, multivariable backward logistic regression was performed and initially included 23 potential factors. The variables with the highest p-value were removed one by one, until remaining variables had $p < 0.157$ (Altman et al., 2009; Royston et al., 2009; Moons et al., 2009a, 2009b). The selection of variables was made over all imputed datasets using Rubin's rules (Wood et al., 2008). To assess whether the level of significance influenced the final prognostic model for all models, selection of

Table 1
Baseline characteristics of 1760 study participants with chronic non-specific low back pain.

Characteristic	Patients ($n = 1760$)	Missing value (n/%)
Number of female patients	1307 (74.3)	0
Age in years: *M* (SD)	40.1 (10.6)	0
Demographic factors		
Low education level	716 (40.7)	71 (4.0)
Marital status/living with one adult	1515 (86.1)	46 (2.6)
Clinical status		
Patients with body mass index > 25	783 (44.5)	88 (5.0)
Duration of complaints in years: *M* (SD)	7.7 (8.8)	0
Cause reported by patient:		23 (1.3)
1 accident/wrong movement	374 (21.3)	
2 after physical overload	73 (4.1)	
3 during pregnancy or after delivery	586 (33.3)	
4 unknown	672 (38.2)	
5 surgery pelvis/back or after HNP	32 (1.8)	
Previous revalidation program	186 (10.6)	101 (5.7)
Co-morbidity	275 (15.6)	88 (5.0)
Pain intensity LBP (VAS in mm): *M* (SD)		
1 present pain intensity	55.5 (23.0)	5 (0.3)
Course of pain intensity due to CNLBP in the previous 3 months		52 (3.0)
1 stable pain intensity	865 (49.1)	
2 increased pain intensity	723 (41.1)	
3 decreased pain intensity	120 (6.8)	
Degree of fatigue LBP (VAS in mm): *M* (SD)		
1 present fatigue	56.5 (26.6)	118 (6.7)
Disability (QBPDS): *M* (SD)	51.7 (15.6)	8 (0.5)
Psychological factors		
Fear avoidance (TSK): *M* (SD)	36.7 (7.3)	50 (2.8)
SCL-90 item 9 *M*(SD)	149.3 (39.7)	227 (12.9)
SF-36 (health-related quality of life)		
PCS	31.8 (7.1)	493 (28.0)
MCS	46.5 (10.3)	493 (28.0)
Work-related factors		
Sickness benefit	924 (52.5)	353 (20.1)
Work participation		161 (9.1)
1 100% working	391 (22.2)	
2 0–99% working	1059 (60.2)	
3 not working*	149 (8.5)	
Physical examination		
ADL function – duration > 31 min without pain increase		
1 walking	410 (23.3)	10 (0.6)
2 sitting	432 (24.5)	13 (0.7)
3 standing	106 (6.1)	9 (0.5)
B200 Isostation (strength) (Newton): *M* (SD)		
1 extension	81.6 (45.8)	107 (6.1)

Values are numbers (percentages) unless stated otherwise in the entire data set of 1760 patients.
M = mean; SD = standard deviation; CNLBP = chronic non-specific low back pain; VAS = Visual analog scale; QBPDS = Quebec Back Pain Disability Scale; TSK = Tampa Scale Kinesiophobia; SCL-90 (item 9) = Symptom Checklist; SF-36 = Short Form; PCS = Physical Component Summary; MCS = Mental Component Summary; ADL = activities of daily living. Missing values ranged from 0.5% ($n = 9$) to 28% ($n = 493$). *"not working" were patients not working at this moment due to seeking new work, or not seeking work because they have family care responsibilities or are retired.

Table 2
Course of quality of life (SF-36) and global perceived effect (GPE) in patients with chronic non-specific low back pain at 2, 5 and 12 months follow-up.

Quality of life (SF-36)	Baseline (n = 1267)	2 months (n = 1252)	5 months (n = 1013)	12 months
PCS; mean (SD)	31.9 (SD 7.1)	40.7 (SD 8.2)	42.1 (SD 10.1)	–
MCS ; mean (SD)	46.6 (SD 10.3)	49.2 (SD 9.4)	50.4 (SD 9.8)	
10% improvement in PCS	–	76.6%	76.3%	–
MCS		39.6%	20.6%	
Global Perceived Effect	Baseline	2 months (n = 981)	5 months (n = 1555)	12 months (n = 976)
1 much improved	–	45.1%	53.0%	60.3%
2 slightly improved		44.1%	32.1%	19.1%
3 no change		7.4%	9.3%	10.8%
4 slightly worsened		3.1%	3.9%	5.7%
5 much worsened		0.3%	1.8%	4.1%
Clinical improvement	–	45.1%	53.0%	60.3%

PCS = Physical Component Scale of the Short Form-36; MCS = Mental Component Scale of the Short-Form 36; mean (SD = standard deviation), n = number of patients.

the variables was repeated with p-values of 0.05. With forward and stepwise selection important variables may be missed in the initial selection phase (Harrell, 2001).

Sensitivity analysis was performed repeating all procedures using GPE as outcome and with a different quality of life cut-off of a 30% improvement on the MCS and PCS with p-values of 0.05 and 0.157 (Ostelo and de Vet, 2005).

2.4.3. Performance of the prognostic model

The performance of the model was checked with regard to the goodness of fit (Hosmer–Lemeshow test), the explained variation, and the discriminative ability of the model. The explained variation of the model is estimated by Nagelkerke's R squared. Explained variation is the extent to which the outcome can be predicted by (the predictors in) the model in current dataset(s). The discriminative ability is reflected by the area under the receiver operating characteristic curve (AUC) [range 0.5 (chance) to 1.0 (perfect discrimination)] (Harrell et al., 1996).

Bootstrapping techniques were used to internally validate the models, i.e. to simulate the performance with respect to the explained variance and the AUC in comparable patient datasets (Vergouwe et al., 2002; Heymans et al., 2007; Moons et al., 2009a, 2009b).

All analyses were done using SPSS version 18.0 (SPSS Inc., USA) and R software (R Foundation for Statistical Computing, Vienna, Austria).

3. Results

3.1. Population

A total of 1760 patients [mean age 40.1 (10.6) years; 74.3% women] with CNSLBP participated in the study. Of these 1760 patients, 1696 (96.4%) completed the 2-month multidisciplinary treatment, 1564 (88.9%) participated in the 5-month follow-up, and 965 (54.8%) completed the 12-month follow-up after start of therapy. Table 1 presents the baseline characteristics of the 1760 patients and the distribution of the possible prognostic factors (Verkerk et al., 2011).

3.2. Course and prognostic models of quality of life

3.2.1. Course at 2 and 5 months

At 2 and 5-months follow-up the mean MCS improved slightly from 46.6 (SD 10.3) at baseline to 49.2 (SD 9.4) at 2 months and to 50.4 (SD 9.8) at 5 months. The mean PCS also improved from 31.9 (SD 7.1) at baseline to 40.7 (SD 8.2) at 2 months and to 42.1 (SD 10.1) at 5 months. At 5 months, a 10% improvement was reported by 20.6% of the patients with regard to the MCS score and by 76.3% with regard to the PCS score (Table 2).

3.2.2. Prognostic factors for improved quality of life at 5-months follow-up

Table 3 shows the associations between potential prognostic factors and PCS and MCS at 5-months follow-up.

The outcome of 10% improvement on the SF-36 PCS was most strongly associated with the following baseline scores: a BMI score $\geq$30 kg/m^2 (OR 1.56, 95% CI 0.96–2.53), receiving sickness benefit (OR 1.90, 95% CI 1.08–3.34), a higher level of work participation (OR 2.03, 95% CI 0.93–4.41), and 16–30 min duration of walking (OR

Table 3
Multivariable models of prognostic factors for 10% improvement in quality of life in patients with chronic non-specific low back pain at 5 months.

Outcome and domains	5-months follow-up		
Outcome Physical Component Scale	OR	95% CI	p-value
Patient characteristics			
Age in years	0.98	0.97–0.99	<0.001
BMI $\geq$25–29.9 kg/m^2	1.14	0.87–1.50	0.334*
BMI $\geq$30 kg/m^2	1.56	0.96–2.53	0.07
Psychological factors			
SF-36 PCS	0.94	0.92–0.96	<0.001
SF-36 MCS	1.03	1.01–1.05	0.01
SCL-90 (item 9)	1.00	0.99–1.01	0.14
Work-related factors			
Sickness benefit (no/yes)	1.90	1.08–3.34	0.03
Work participation	2.03	0.93–4.41	0.07
Physical examination			
Duration walking 1 (0–15 min)	1.19	0.75–1.89	0.419*
Duration walking 2 (16–30 min)	1.78	1.08–2.97	0.03
Duration walking 3 (31–45 min)	1.68	0.77–3.69	0.17*
Outcome Mental Component Scale	OR	95% CI	p-value
Patient characteristics			
Gender (female/male)	0.70	0.43–1.13	0.13
Psychological factors			
SF-36 PCS	1.03	1.00–1.07	0.05
SF-36 MCS	0.82	0.79–0.84	<0.001
SCL-90 (item 9)	0.99	0.99–1.00	<0.001

95% CI = 95% confidence interval, OR = odds ratio, an OR > 1 reflects a higher probability of 10% recovery for the outcome PCS and MCS and an OR < 1 a lower probability of 10% recovery for the outcome back pain intensity, compared to the reference category. OR estimated after multiple imputation (n = 5 datasets) with p-value of 0.157.
SCL-90 (item 9) = Symptom Checklist; SF-36 = Short Form; PCS = Physical Component Summary; MCS = Mental Component Summary. The variable body mass index (BMI) is a category value of 3 (18–24.9 kg/m^2; $\geq$25–29.9 kg/m^2; $\geq$30 kg/m^2); the variable duration walking is a category value of 4 (0–15; 16–30; 31–45; >60).

1.78, 95% CI 1.08–2.97). The AUC of this model was 0.69 and the explained variance was 11%.

The factors most strongly associated with a 10% improvement on the MCS score were being female (OR 0.70, 95% CI 0.43–1.113) and having a lower MCS score at baseline (OR 0.82, 95% CI 0.79–0.84). The AUC of this model was 0.88 and the explained variance was 44%.

With regard to internal validation of the models, for PCS the explained variance at 5-months follow-up was 12% with an AUC of 0.69; for MCS these figures were 44% and 0.88, respectively.

3.3. Course and prognostic models of GPE

3.3.1. Course at 5 and 12-months follow-up

At 5 and 12-months follow-up, clinical improvement was reported by 53% and 60.3% of the patients, respectively. In addition, at 5 and 12-months follow-up, 'no change too much worsened' was reported by 15% and 20.6% of the patients, respectively (Table 2).

3.3.2. Prognostic factors for GPE at 5 and 12-months follow-up

Table 4 shows associations between potential prognostic from the predefined time point (i.e. after 2 months of therapy at SJC) and GPE of the patients at 5 and 12-months follow-up.

Being married (OR 1.39, 95% CI 1.00–1.91), decrease of pain intensity in the last 3 months (OR 2.07, 95% CI 1.23–3.48), receiving sickness benefit (OR 1.61, 95% CI 0.96–2.69) and a higher work participation (OR 1.92, 95% CI 1.03–3.59) were the strongest factors associated with clinical improvement on the GPE scale at 5-months follow-up. At 12-months the following factors showed the strongest associations: being female (OR 0.63, 95% CI 0.47–0.84), being married (OR 1.51, 95% CI 1.03–2.21), higher work participation at baseline (OR 1.65, 95% CI 1.18–2.29) and duration of walking 16–31 min at baseline (OR 1.58, 95% CI 0.88–2.82).

The explained variance and AUC for 5 and 12-months were 11% and 0.66, and 9% and 0.65, respectively. The internal validation showed similar results in the GPE for explained variance and AUC.

3.4. Sensitivity analysis regarding quality of life and GPE

Repeating the analysis with *p*-values of 0.05 or 0.157, and using CCA or 5 or 40 imputated datasets or a different quality of life cut-off of a 30% improvement on the MCS and PCS, resulted in similar prognostic factors for a 10% improvement in the PCS, MCS and GPE-score at 5-months follow-up. At 12 months, younger follow-up age, less pain intensity at baseline, higher work participation or shorter duration of complaints were often related to GPE in the different models. The explained variance, AUC and internal validation were similar to earlier findings.

4. Discussion

In the present study, a main finding is the sustained 10% improvement on the PCS (76.3% of the population) up to 5 months and on GPE (60.3%) up to 12 months. For MCS this 10% improvement is slightly less (20.6%) at 5 months, but a mean of 50 (SD of 10) represents normal health and function (Walsh et al., 2003). Some patients reported no improvement on GPE at 5 and 12-months follow-up (15% and 20.6%, respectively).

The present study shows that improvement in quality of life (on SF-36 MCS) at 5-months follow-up was associated with patients' characteristics and psychological factors. At 5-months, improvement on quality of life (on SF-36 PCS) and GPE was associated with patients' characteristics, physical examination, work-related factors and psychological factors. For GPE, clinical status was also associated with improvement.

At 12-months follow-up GPE was associated with patients' characteristics, clinical status, physical examination and work-related factors. The sensitivity analyses showed overall similarity for the prognostic factors. The prognostic models provide

Table 4
Multivariable models of prognostic factors for absolute recovery in chronic non-specific low back pain, global perceived effect (GPE) at 5 and 12 months follow-up.

Outcome and domains	5-months follow-up			12-months follow-up		
Outcome GPE	OR	95% CI	*p*-value	OR	95% CI	*p*-value
Patient characteristics						
Age in years	0.97	0.96–0.99	<0.001	0.98	0.97–0.99	0.002
Gender (female/male)				0.63	0.47–0.84	0.002
Married/being with one adult (no/yes)	1.39	1.00–1.91	0.05	1.51	1.03–2.21	0.03
Clinical status						
Duration of complaints	0.99	0.97–1.00	0.05	0.98	0.97–0.99	0.02
Course of pain intensity due to CNLBP in the previous 3 months (increase of pain)	1.05	0.84–1.30	0.681*			
Course of pain intensity due to CNLBP in the previous 3 months (decrease of pain)	2.07	1.23–3.48	0.007			
Back pain intensity (VAS)	1.00	0.99–1.00	0.09	0.99	0.98–0.99	<0.001
Disability (QBPDS)	1.01	1.00–1.02	0.01			
Psychological factors						
TSK	0.97	0.96–0.99	0.005			
SF-36 PCS	1.05	1.03–1.07	<0.001			
Sf-36 MCS	1.02	1.00–1.03	<0.001			
Work-related factors						
Sickness benefits (no/yes)	1.61	0.96–2.69	0.07			
Work participation	1.92	1.03–3.59	0.04	1.65	1.18–2.29	0.005
Physical Examination						
B200 Isostation extension	1.00	0.99–1.00	0.08			
Duration walking 1 (0–15 min)				1.00	0.74–1.36	0.99*
Duration walking 2 (16–30 min)				1.58	0.88–2.82	0.11
Duration walking 3 (31–45 min)				1.32	0.89–1.96	0.16

95%-CI = 95% confidence interval, OR = odds ratio, OR estimated after multiple imputation ($n = 5$ datasets) with *p*-value of 0.157.
VAS = Visual Analog Scale; QBPDS = Quebec Pain Disability Scale; SF-36 = Short Form; PCS = Physical Component Summary; MCS = Mental Component Summary. The variable course of pain intensity due to CNLBP in the previous 3 months is a category variable, 1) increase, 2) decrease and 3) stable of pain intensity; the variable duration walking is a category value of 4 (0–15; 16–30; 31–45; >60).

additional information to present a more realistic expectation regarding outcome. However, development of a prognostic model does not involve investigating the causal associations between individual prognostic factors and outcome.

Comparison of the present results with earlier studies is limited, because our systematic review identified only 3 low-quality studies addressing this topic (Verkerk et al., 2012). For the outcome SF-36 PCS, Keeley et al. (2008) had a 6-months follow-up with a mean SF-36 PCS score of 34.9 (SD 10.9) compared with a baseline score of 33.3 (SD 10.; $n = 93$). The present study showed a greater improvement on the PCS score, i.e. from 31.9 (SD 7.1) at baseline to 42.1 (SD 10.1) at 5-months. The difference in results might be due to differences in study characteristics; e.g. patients in the study of Keeley et al. (2008) did not follow a therapy program but could contact their healthcare provider when needed; the authors concluded that an intervention targeting these psychosocial variables in patients, may lead to improved quality of life and reduction of healthcare costs (Keeley et al., 2008). A study by van der Hulst et al. (2008) from the same systematic review on SF-36 PCS and MCS, showed more similarity with the present study at 6-months follow-up. Their patients with CNLBP experienced (on average) better health-related quality of life than at baseline, regardless of the type of treatment [Roessingh Back Rehabilitation Program (RBRP) vs. usual care]. At follow-up the RBRP (7-week program) resulted in a PCS score of 37 (SD 9) and a MCS score of 51 (SD 9) compared with baseline scores of 31 (SD 7) and 49 (SD 10), respectively (van der Hulst et al., 2008).

In relationship to the course of GPE, our systematic review (Verkerk et al., 2012) found only one study, which reported that 29% of the non-surgical group assessed themselves as improved at 2-year follow-up (Hagg et al., 2003a). In contrast, 60.3% of our patients reported clinical improvement on the GPE scale at 12-months follow-up.

In the final prognostic model on PCS reported by Keeley et al. (2008), the Hospital Anxiety and Depression Scale (HADS) total score and back-pain related social stress, continued to make a significant contribution to the model ($R^2 = 0.72$; incidence rate ratio around 1.00). In the present study, the psychological factors [SCL-90 (item 9) and MCS; OR around 1.00] were included, as were other factors with a strong association. In both studies, the psycho-social results had a low association; further research on these items is necessary.

In the present study no association was found for the factor 'fear avoidance beliefs' and the outcome PCS. In two of the studies in our systematic review (Verkerk et al., 2012) conflicting evidence was found for their 8-week (van der Hulst et al., 2008) and 6-month (Keeley et al., 2008; van der Hulst et al., 2008) follow-up, whereas the 6-month follow-up data of Keeley et al. (2008) are similar to those in the present study. The discrepancy between these results may be due to differences in characteristics between the two studies, including a smaller patient population ($n < 200$), differences concerning treatment/no treatment, in the length of follow-up (8 weeks), and in the included prognostic factors. In the 2-months therapy at the SJC and in the 7-week RBRP program of van der Hulst et al. (2008), fear of avoidance beliefs was a part of the program but yielded differing results, possibly due to other aspects of the therapy program. In van der Hulst's study (2008), presence at work predicted improvement for the PCS at 6-months follow-up, which is in line with our results at 5-months follow-up. This might be explained by the fact that people at work are generally healthier and more physically active, which may be related to greater physical wellbeing. However, because this comparison is with only one study, more research is needed on this topic. Also, in van der Hulst's study (2008), whereas higher depression scores (SCL-90-dep) predicted deterioration on the MCS on the short and long-term follow-up regardless of treatment, this was in contrast to our results.

For the outcome GPE only one study was found (Verkerk et al., 2012), reporting that increased pre-treatment depressive symptoms measured with the Zagazig Depression Scale predicted improvement of the GPE score in a non-surgical group of CNSLBP patients (Hagg et al., 2003a), We found no association with the SCL-90 (item 9) in the final model at 5 and 12-months follow-up.

The outcome quality of life and GPE were two of the 5 outcomes (back pain intensity, disability due back pain, work participation, quality of life and patients' perceived recovery) measured in this prospective cohort study. The choice for these current outcomes is because this is important to the patient and clinician. Also only a few other studies are known about this topic.

Of all patients, 90.2% had stable or increased low back pain intensity in the 3 months prior to intake (Verkerk et al., 2011). The duration of complaints in our study population was on average 7.7 years. During the 12 months there we those patients that recovered from back pain, those who experience it off and on and those who have it most of the time. (Axen and Leboeuf-Yde, 2013; Dunn et al., 2013a, 2013b) Recent studies (Axen and Leboeuf-Yde, 2013; Dunn et al., 2013a) report that most patients with back pain appear to follow a particular pain trajectory over longer time periods, and do not have frequently recurring of widely fluctuating patterns. It can be that a particular pain trajectory will have certain clinical characteristics. This could influence which prognostic factor is important as also the effect in rehabilitation (Axen and Leboeuf-Yde, 2013).

The present study also has some limitations. First, despite the large sample size, at baseline there were missing values (0.5–28%). Also, at 12-months follow-up only 54.8% of the patients could be compared with the baseline measurements. Our study gathered the data at the rehabilitation centre SJC during daily clinical practice and at 12 months this was done by postal questionnaire. The SF-36 (28% missing values at the baseline) is collected electronically and separately from the other data at the start of therapy by a therapist. The general practitioner (GP) asked at baseline which kind of sickness benefit (20%) a patient had. Sometimes this was forgotten. Other reasons for missing data could be that the patient didn't understood all the questions or an incomplete questionnaire was retrieved. Loss to follow-up (i.e. failure to return the follow-up questionnaires) occurred for various reasons, including vacation, envelope not stamped, recovered from CNSLBP, did not find it necessary, starting another intervention, etc. No reminder was sent to the patient, this was not a part of the daily clinical care at the SJC.

We assume that the missing values occurred at random, which is not uncommon with a long-term follow-up. Also, we used imputation of data (multiple imputation techniques); however, this is reported to be a valid method to deal with missing values (van Buuren, 2012) and the sensitivity analyses yielded similar results.

We cannot demonstrate the influence of the given cognitive behaviour therapy with supervised exercises, educational and multi-disciplinary treatment. Only, that this is one of the most common intervention for CNSLBP in Dutch rehabilitation centres and two Cochrane reviews (Guzmán et al., 2002; Henschke et al., 2010) provided evidence of a greater improvement on the short term than other treatments.

For the present study, although we chose for a cut-off point of 10% improvement on the SF-36 PCS and MCS, there was little difference in identifying prognostic factors when a 30% improvement was used. A 30% improvement is a more commonly used criteria in CNSLBP especially for the outcome pain and disability (Ostelo and de Vet, 2005; Ostelo et al., 2008); however, the problem remains that patients close to, but on opposite sides of the cut-off point, are characterized as being very different rather than very similar. Also,

although the currently available GPE scale has the option 'completely recovered' (Jellema et al., 2005), this was not yet in use in the SJC when the data were retrieved. This latter outcome measure is often dichotomized because it is easier for interpretation by clinicians and patients, albeit with the risk of losing some information (Altman and Royston, 2006). Because patients have difficulty taking their baseline status into account when scoring the GPE scale (Kamper et al., 2010), this item was compared with the end of therapy at the SJC.

Further research should focus on (external) validation of the presented prognostic models with appropriate study methodology, rather than developing new ones. With further testing the practical value of the models can be properly established (Moons et al., 2009a). The impact for the clinician is that the current thought suggests a more complex interaction between factors rather than singular prognostic factors that influence the patient through time. There is clearly a need to investigate how prognostic factors work together in their usefulness and feasibility in clinical practice.

Acknowledgment

The study was financial supported by Rotterdam University of Applied Sciences and Department of General Practice, Erasmus MC. The authors declare no conflict of interest.

References

Aktekin LA, Eser F, Malhan S, Öksüz E, Keskin D, Bodur H. A comparison of four different HRQoL generic questionnaire in five different patient groups. Rheumatol Int 2009;30:63–7.
Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ 2006;332:1080.
Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ 2009;338(b605):1432–5.
Axen I, Leboeuf-Yde C. Trajectories of low back pain. Best Pract Res Clin Rheumatol 2013;27:601–12.
Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. Spine (Phila Pa 1976) 2000;25:3100–3.
Costa Lda C, Maher CG, McAuley JH, Hancock MJ, Herbert RD, Refshauge KM, et al. Prognosis for patients with chronic low back pain: inception cohort study. BMJ 2009;339(b3829).
Davidson M, Keating JL, Eyres S. A low back-specific version of the SF-36 Physical Functioning scale. Spine (Phila Pa 1976) 2004;29:586–94.
Dunn KM, Campbell P, Jordan KP. Long-term trajectories of back pain: cohort study with 7-year follow-up. BMJ Open 2013a;3:e003838.
Dunn KM, Hestbaek L, Cassidy JD. Low back pain across life course. Best Pract Res Clin Rheumatol 2013b;27:591–600.
Gandek B, Sinclair SJ, Kosinski M, Ware Jr JE. Psychometric evaluation of the SF-36 health survey in Medicare managed care. Health Care Financ Rev 2004;25:5–25.
Gandek B, Ware Jr JE, Aaronson NK, Alonso J, Apolone G, Bjorner J, et al. Tests of data quality, scaling assumptions, and reliability of the SF-36 in eleven countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol 1998;51:1149–58.
Guzmán J, Esmail R, Karjalainen K, Malmivaara A, Irvin E, Bombardier C. Multidisciplinary bio-psycho-social rehabilitation for chronic low back pain. Cochrane Database Syst Rev 2002;(1):CD000963.
Hagg O, Fritzell P, Ekselius L, Nordwall A. Predictors of outcome in fusion surgery for chronic low back pain. A report from the Swedish Lumbar Spine Study. Eur Spine J 2003a;12:22–33.
Hagg O, Fritzell P, Nordwall A. The clinical importance of changes in outcome scores after treatment for chronic low back pain. Eur Spine J 2003b;12:12–20.
Harrell Jr FE. Regression modelling strategies. New York, NY: Springer; 2001.
Harrell Jr FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361–87.
Heneweer H, Aufdemkampe G, van Tulder MW, Kiers H, Stappaerts KH, Vanhees L. Psychosocial variables in patients with (sub)acute low back pain: an inception cohort in primary care physical therapy in The Netherlands. Spine (Phila Pa 1976) 2007;32:586–92.
Henschke N, Ostelo RW, van Tulder MW, Vlaeyen JW, Morley S, Assendelft WJ, et al. Behavioural treatment for chronic low-back pain. Cochrane Database Syst Rev 2010;7:CD002014.
Heymans MW, van Buuren S, Knol DL, van Mechelen W, de Vet HC. Variable selection under multiple imputation using the bootstrap in a prognostic study. BMC Med Res Methodol 2007;7:33.
Jellema P, van der Windt DAWM, van der Horst HE, Twisk JWR, Stalman WAB, Bouter LM. Should treatment of (sub)acute low back pain be aimed at psychosocial prognostic factors? Cluster randomised clinical trial in general practice. BMJ 2005:1–7.
Kamper SJ, Ostelo RW, Knol DL, Maher CG, de Vet HC, Hancock MJ. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. J Clin Epidemiol 2010;63:760–7666.e1.
Keeley P, Creed F, Tomenson B, Todd C, Borglin G, Dickens C. Psychosocial predictors of health-related quality of life and health service utilisation in people with chronic low back pain. Pain 2008;135:142–50.
Kent PM, Keating JL. Can we predict poor recovery from recent-onset nonspecific low back pain? A systematic review. Man Ther 2008;13:12–28.
Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009a;338(b606):1487–90.
Moons KG, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: what, why, and how? BMJ 2009b;338(b375):1317–20.
Ostelo RW, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine (Phila Pa 1976) 2008;33:90–4.
Ostelo RWG, de Vet HCW. Clinically important outcomes in low back pain. Best Pract Res Clin Rheumatol 2005;19:593–607.
Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49:1373–9.
Pincus T, Santos R, Breen A, Burton AK, Underwood M. A review and proposal for a core set of factors for prospective cohorts in low back pain: a consensus statement. Arthritis Rheum 2008;59:14–24. http://dx.doi.org/10.1002/art.23251.
Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: developing a prognostic model. BMJ 2009;338(b604):1373–7.
Snyder-Halpern R. Indicators of organizational readiness for clinical information technology/systems innovation: a Delphi study. Int J Med Inform 2001;63:179–204.
Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. Stat Med 2004;23:2567–86.
van Buuren S. Flexible imputation of missing data. Boca Raton: Chapman & Hall/CRC; 2012.
van der Hulst M, Vollenbroek-Hutten MM, Groothuis-Oudshoorn KG, Hermens HJ. Multidisciplinary rehabilitation treatment of patients with chronic low back pain: a prognostic model for its outcome. Clin J Pain 2008;24:421–30.
van Oort L, van den Berg T, Koes BW, de Vet RHCW, Anemad HJR, Heymans MW, et al. Preliminary state of development of prediction models for primary care physical therapy: a systematic review. J Clin Epidemiol 2012;65:1257–66.
Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Validity of prognostic models: when is a model clinically useful? Semin Urol Oncol 2002;20:96–107.
Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Bouter LM, et al. The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. J Clin Epidemiol 1998;51:1235–41.
Verkerk K, Luijsterburg PA, Ronchetti I, Miedema HS, Pool-Goudzwaard A, van Wingerden JP, et al. Course and prognosis of recovery for chronic non-specific low back pain: design, therapy program and baseline data of a prospective cohort study. BMC Musculoskelet Disord 2011;12:252.
Verkerk K, Luijsterburg PA, Miedema HS, Pool-Goudzwaard A, Koes BW. Prognostic factors for recovery in chronic nonspecific low back pain: a systematic review. Phys Ther 2012;92:1093–108.
Walsh TL, Hanscom B, Lurie JD, Weinstein JN. Is a condition-specific instrument for patients with low back pain/leg symptoms really necessary? The responsiveness of the Oswestry Disability Index, MODEMS, and the SF-36. Spine (Phila Pa 1976) 2003;28:607–15.
Wood AM, White IR, Royston P. How should variable selection be performed with multiply imputed data? Stat Med 2008;27:3227–46.