Residual complaints following lumbar disc surgery: prognostic indicators of outcome

Pain 114 (2005) 177–185

PAIN

www.elsevier.com/locate/pain

Residual complaints following lumbar disc surgery: prognostic indicators of outcome

Raymond W.J.G. Ostelo*, Johan W.S. Vlaeyen, Piet A. van den Brandt, Henrica C.W. de Vet

VU University Medical Center, EMGO Institute, van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands

Received 17 June 2004; received in revised form 26 November 2004; accepted 13 December 2004

Abstract

Physical as well as psychological features might be important prognostic factors for residual complaints following lumbar disc surgery in primary care. No studies have yet investigated both factors simultaneously. The aim of this prospective cohort study was to identify indicators of the short and long-term outcome of residual complaints following lumbar disc surgery. Patients ($n=105$), aged between 18 and 65 years, were included if they still suffered residual complaints 6 weeks after first-time lumbar disc surgery and had therefore been referred to physiotherapy. All potential indicators were measured at baseline except treatment expectancy, which was measured after two treatment sessions enabling patients to rate treatment expectancy based on their actual perception of the treatment. Dimensions of recovery included perceived recovery, functional status, and pain intensity (back and leg) at the 3-month and 12-month follow-up. It was found that high treatment expectancy was associated with a favorable outcome on perceived recovery and functional status, both at the 3 and the 12-month follow-up. Taking pain medication and a poor functional status at baseline were associated with poor perceived recovery and functional status at both follow-up measurements. Leg pain and back pain at baseline were associated with residual leg and back pain at the 3 and the 12-month follow-up, respectively. The results for perceived recovery and functional status were rather robust. However, for leg pain and back pain, the results were less stable. Apparently, the clinical course to recovery of residual leg pain and residual back pain is not strongly influenced by these indicators.

Keywords: Prospective study; Prognosis; Lumbar disc surgery; Low back pain; Treatment expectancy

1. Introduction

The occurrence of residual complaints following lumbar disc surgery ranges from 22% to 45% and 30% to 70% of patients report residual sciatica and residual low back pain, respectively, after surgery (Dvorak et al., 1988; Hurme and Alaranta, 1987; Korres et al., 1992; Manniche et al., 1994a, b; Pappas et al., 1992; Weber, 1983; Yorimitsu et al., 2001). The duration and intensity of the post-surgery rehabilitation varies widely, but intensive exercise programs seem to be the most effective (Ostelo et al., 2003b). The rehabilitation of these patients takes place predominantly in a primary care setting. Identifying prognostic factors that predict the clinical course of residual complaints might be important for the further development of effective methods of treatment, especially when these prognostic factors can be modified.

In primary care, some prognostic factors are routinely included in history-taking, for example age, gender, and severity of pain. Physical features such as height and weight, and the activities of daily life (ADL) have been reported to be important in predicting improvement in functional status (Hurme and Alaranta, 1987). Additionally, psychological factors might also influence the clinical course of residual complaints. Kjelby-Wendt (Kjellby- Wendt et al., 1999) reported that patients who became more discontented 3 and 12 months after the surgery were more depressed, more anxious, and had experienced more pain before the surgery. Psychological factors, in particular, could be influenced by the fact that a patient underwent surgery.

* Corresponding author. Tel.: +31 20 4448149; fax: +31 20 4448181.
E-mail address: r.ostelo@vumc.nl (R.W.J.G. Ostelo).

0304-3959/$20.00
doi:10.1016/j.pain.2004.12.021

To our knowledge, no studies have yet assessed the influence of these psychological factors on post-surgery measurements. In analogy with the etiology of chronic low back pain (Fritz et al., 2001; Vlaeyen and Linton, 2000), it can be hypothesized that an excessively negative orientation toward pain (pain catastrophizing) and fear of movement/(re-)injury (kinesiophobia) after the surgery have an important influence on the development or maintenance of residual complaints following surgery. The mechanism has been described as follows (Picavet et al., 2002): "Persons who catastrophically misinterpret innocuous bodily sensations, including pain, are likely to become fearful of pain, which results in at least two processes. First, pain-related fear is associated with avoidance behaviors and the avoidance of movement and physical activity in particular. Second, pain-related fear is associated with increased bodily awareness and pain hypervigilance. Hypervigilance, depression, and disuse are known to be associated with increased pain levels and hence might exacerbate the painful experience".

The patient's treatment expectations also seem to be important in predicting outcomes in low back pain (Kalauokalani et al., 2001) as well as the outcomes of surgery for sciatica (Lutz et al., 1999). However, it is unknown whether fear-avoidance beliefs, kinesiophobia and the patient's treatment expectations are factors that influence the clinical course following lumbar disc surgery. Moreover, as far as we know, no studies have yet investigated physical and psychological factors simultaneously. Primary care physicians (e.g. physiotherapists, general practitioners) can relatively easily assess these variables by history-taking or by the use of short questionnaires.

The aim of this prospective study is to investigate whether the findings of history-taking, physical factors, psychological assessment and the patient's treatment expectations predict the course of residual complaints during rehabilitation following lumbar disc surgery.

2. Materials and methods

2.1. Design

Our prospective study was conducted within the framework of a randomized controlled trial on the effectiveness of usual care provided by physiotherapists, compared to a behavioral graded activity (BGA) program (Ostelo et al., 2003a). Both methods of treatment, delivered in primary care, consisted of 18, 30-min sessions within 3 months. Usual physiotherapeutic care consisted of exercises to increase muscle strength, stability and mobility in order to improve ADL. The patients were also instructed how to lift, sit, and stand in the right way and how to perform various kinds of ADL. If necessary, the therapists were allowed to apply electrotherapy, massage or manipulations. The BGA program, consisted of operant treatment, using graded activity and positive reinforcement to increase health behaviors and decrease pain behaviors, based on time-contingency management. The essence of the BGA was to establish an individually tailored exercise program with increasing intensity, based on the measurements on intake. The physical therapists underwent a 2-day training session, and they received follow-up training during the course of the trial. In the current study these interventions were not considered to be prognostic indicators, but were adjusted for in all analyses.

2.2. Subjects

The neurosurgeon in each of the four participating hospitals in the south of The Netherlands recruited the patients during the routine from 6-weeks post-surgery visits. Patients (aged between 18 and 65 years) were included at the time of this visit if they still suffered from residual complaints (severe back or leg pain and/or restrictions in ADL and/or work) that justified a referral to physiotherapy. This resembles the standard procedures adhered to in the participating hospitals. Patients were excluded if there had been complications during surgery, if they had a confirmed and relevant underlying disease (e.g. stenosis or M. Bechterew), or if one of the treatments was contraindicated. The Medical Ethics Committee of the University Hospital Maastricht (The Netherlands) approved the study protocol.

2.3. Baseline measurements

At baseline, i.e. before the first treatment session the following potential prognostic indicators were measured: gender, age, duration of complaints before surgery, whether or not (pain)medication was taken at baseline because of the residual complaints, number of days in hospital following the surgery, body mass index (BMI), severity of pain in back and leg (both on a VAS), pain catastrophizing (Pain Catastrophizing Scale, PCS) (Sullivan et al., 1995), fear of movement (Tampa Scale for Kinesiophobia, TSK) (Kori et al., 1990) and negative affectivity (Negative Emotionality sub-scale of the Multidimensional Personality Questionnaire) (Stegen et al., 1998; Tellegen, 1982). Finally, at baseline, the patients rated their confidence in recovery in general ('great deal', 'moderate', 'no confidence' and 'do not know'). This last variable seems to have high face validity, and was therefore included to assess its predictive power. The patient's treatment expectancy (11-point Likert scale: 0 = expects no benefit at all, 10 = absolutely convinced of benefit) was the only variable that was not measured at baseline but after two treatment sessions. The reason for this was to enable patients to rate their expectations based on their actual perception of the treatment, and not simply their assumptions (Vlaeyen et al. 1996).

2.4. Dimensions of recovery

For patients with residual complaints following lumbar disc surgery, several dimensions of recovery are frequently reported: perceived recovery, functional status, and pain intensity (back and leg).

Therefore, the following questionnaires, completed by the patients themselves, were administered during the 3 (post-treatment) and 12-month follow-up measurements.

(1) Patients rated their perceived recovery on a 7-point ordinal transition scale, ranging from 'completely recovered' to

'worse than ever'. A priori recovery was defined as 'completely recovered' or 'much improved', as reported by the patient.

(2) The Roland-Disability Questionnaire (RDQ) measured functional status (Roland and Morris, 1983). A priori recovery was defined as an RDQ score of ≤ 4. As this is a somewhat arbitrary cut-off score, we performed sensitivity analyses with different RDQ cut-off values (≤ 3 and ≤ 5), in order to assess the robustness of the findings.

(3) Severity of back and leg pain (average pain previous week) was scored on a VAS. A priori recovery was defined as a VAS score of ≤ 10 mm. Again, a sensitivity analysis was performed, using a cut-off value of ≤ 20 mm, to assess the robustness of the findings.

2.5. Statistical analysis

First, the relationship between each potential prognostic indicator and the outcome at issue was individually evaluated for the short-term and the long-term separately, using logistic regression and only adjusting for intervention. Odds Ratios (ORs) and 95% confidence interval (CI) were calculated for each dimension of recovery, to reflect the strength of the association. Variables showing a promising relationship (P-value ≤ 0.25) were included in the multivariate logistic regression. Subsequently a backward elimination (P-value ≤ 0.10) was applied retaining only those variables that were strongly associated with the outcome at issue. The proportion of explained variance (Nagelkerke R^2) and the percentage of correctly classified patients are presented to give an indication of the predictive power of the model. A patient was considered to be classified correctly if the predicted probability (according to the prognostic model) of being in the observed group (recovered versus not recovered for the outcome measure at issue) was more than 0.5. Occasional missing values in the prognostic indicators at baseline ($<5\%$) were replaced by their group means.

If patients did not attend a follow-up measurement, or dropped out of the study completely, the reason for this was taken into account in the replacement of the missing values. We think that this procedure best anticipates the information at our disposal. The reasons for not attending the measurement were registered by the research assistant, who was blinded for the treatment allocation and performed all measurements. Then three senior epidemiologists, also blinded for the treatment allocation, evaluated these reasons per patient independently, based on pre-set criteria, and decided whether the patient should be considered to have 'recovered' or 'deteriorated'. Patients were considered to have 'recovered' if they could not attend a measurement because, for example, they had returned to work completely and were therefore no longer willing (or able) to attend the measurements. Patients were considered to have 'deteriorated' if they did not attend a measurement because they had more pain and consequently no longer wished to participate in the study. Patients were also considered to have 'deteriorated' if the neurosurgeon advised that the treatment should be stopped because of strong indications of a (new) herniated disc. If patients could not be considered to have 'improved' or 'deteriorated', and if there was no association with the allocated treatment (e.g. the patient moved out of the catchment area), missing values were replaced by mean scores. For the substitution of 'improved' or 'deteriorated' we used the 10th or 90th percentile score of the total group. For example, the 10th percentile was used to substitute low RDQ because these low RDQ scores indicate a low level of disability, and the 90th percentile was used to substitute the SF-36 sub-scales because high scores indicate a good health status.

3. Results

From November 1997 until December 1999 a total of 671 patients were checked for eligibility during their routine 6 weeks post-surgery visit to the neurosurgeon. Fifty-seven percent of the patients had more or less recovered (no further symptoms). One hundred and forty one patients (21%) were excluded for various reasons: although operated in a participating hospital, not living in the catchment area ($n=40$), not motivated to participate ($n=32$), too old ($n=30$), presented with co-morbidities ($n=22$), language problems ($n=8$), previous lumbar surgery ($n=7$), and insurance problems ($n=2$). We failed to trace 43 patients (6%). In total, 105 (16%) patients had been referred to physiotherapy because of the residual complaints and were eligible for participation. After receiving all information regarding the study the participants signed informed consent. At the 3-month measurement, eight patients dropped out. The replacement procedure yielded three mean substitutions, three positive substitutions, and two negative substitutions. At the 12-month follow-up, another four patients dropped out; two negative and two mean substitutions. Table 1 presents the number of patients who were considered to have recovered at 3 and 12 months.

3.1. Perceived recovery

Table 2 presents the prognostic indicators of perceived recovery.

Taking pain medication, poor functional status, high levels of negative affectivity and low treatment expectancy were associated with poor perceived recovery at the 3-month

Table 1
Number of patients who had recovered at 3 and 12 months and the various dimensions of recovery

Outcomes	Recovered (n)	Success rate (%)
Perceived recovery		
3 Months	59	56.2
12 Months	74	70.5
Functional status (RDQ-score: ≤ 4)		
3 Months	39	37.1
12 Months	46	43.8
Leg pain[a] (VAS-score: ≤ 10 mm)		
3 Months	43	41.0
12 Months	52	49.5
Back pain[a] (VAS-score: ≤ 10 mm)		
3 Months	31	29.5
12 Months	46	43.8

RDQ, Roland Disability Questionnaire; VAS, Visual Analogue Scale.
[a] Average pain previous week.

Table 2
Prognostic indicators of perceived recovery

Variable	3 Months				12 Months			
	Individual indicators		Multiple regression		Individual indicators		Multiple regression	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Gender (female) ($n=45$)	1.28	0.57; 2.87			0.50	0.21; 1.19*		
Age (≥ 43)[a]	0.49	0.21; 1.10*			0.57	0.24; 1.35*		
Duration of complaints before surgery								
(1) ≥ 2–≤ 6 months ($n=56$)	1.00				1.00			
(2) ≥ 7–≤ 12 months ($n=40$)	1.15	0.49; 2.66			0.97	0.39; 2.41		
(3) >12 months ($n=9$)	0.38	0.08; 1.74*			0.30	0.07; 1.24*		
Medication at baseline (yes) ($n=38$)	0.49	0.21; 1.11*	0.52	0.24; 1.14	0.40	0.17; 0.94*	0.45	0.20; 0.99
Days in hospital (>7 days) ($n=28$)	0.76	0.31; 1.87			1.75	0.63; 4.88		
Body mass index								
(1) <25 ($n=43$)	1.00				1.00			
(2) ≥ 25, <30 ($n=47$)	0.59	0.25; 1.41			1.26	0.51; 3.11		
(3) ≥ 30 ($n=15$)	0.64	0.19; 2.16			1.32	0.36; 4.90		
Functional status (RDQ >13)[a]	0.36	0.15; 0.83*	0.46	0.21; 1.00	0.30	0.11; 0.80*	0.40	0.17; 0.96
Pain back (VAS >45)[a]	0.66	0.30; 1.47			1.02	0.44; 2.38		
Pan leg (VAS >43)[a]	0.56	0.25; 1.24*			0.73	0.32; 1.70		
Fear of movement (TSK >40) ($n=77$)	1.26	0.51; 3.09			1.10	0.41; 2.75		
Catastrophizing (PCS)								
(1) ≤ 11	1.00				1.00			
(2) >11–≤ 21	0.62	0.23; 1.64			0.90	0.30; 2.75		
(3) >21	0.51	0.20; 1.38*			0.35	0.12; 0.98*		
Negative affectivity (NEM)								
(1) ≤ 1	1.00				1.00			
(2) >1–≤ 4	0.63	0.24; 1.66	0.63	0.26; 1.55	0.62	0.22; 1.78		
(3) >4	0.34	0.27; 0.94*	0.38	0.16; 0.93	0.46	0.16; 1.34		
Confidence in recovery in general								
(1) Great deal ($n=43$)	1.00				1.00			
(2) Moderate ($n=42$)	1.63	0.65; 4.09			1.26	0.48; 3.36		
(3) No confidence ($n=6$)	0.13	0.01; 1.26*			0.39	0.07; 2.20		
(4) Do not know ($n=14$)	0.30	0.08; 1.11*			0.49	0.14; 1.74		
Patients' treatment expectancy (≥ 6) ($n=80$)	3.78	1.40; 10.21*	3.11	1.30; 7.48	3.54	1.27; 9.88*	2.52	1.09; 5.87
Nagelkerke R^2			0.233				0.151	
Correctly classified (0.500)			69.5%				69.5%	

Odds Ratios (OR) and 95% Confidence Interval (CI) for each indicator. All analyses adjusted for intervention. *P-value ≤ 0.250 for individual indicator, included in multivariate logistic regression. Variables with more than two categories were included if only one category had a P-value ≤ 0.250. In multivariate logistic regression only variables with P-value ≤ 0.10. RDQ, Roland Disability Questionnaire; VAS, Visual Analogue Scale; TSK, Tampa Scale for Kinesiophobia; PCS, Pain Catastrophizing Scale, NEM, Negative Emotionality subscale.

[a] Cut-off value = median, for all other variables numbers in categories between brackets. For dichotomous variables the reference category is the contrast (female versus male), for all other variables the reference category is first category (OR = 1).

follow up. The results for the 12-month follow-up were rather similar: only negative affectivity was no longer associated with perceived recovery. The overall predictive power of the models was only modest; Nagelkerke R^2 was 0.233 and 0.151 for 3 and 12 months, respectively. The percentage of correctly classified patients was 69.5% for both follow-up measurements.

3.2. Functional status

Table 3 presents the prognostic indicators of functional status.

Taking pain medication, a poor functional status, a high BMI and a low treatment expectancy were associated with poor recovery of functional status at the 3-month follow-up. The results for the 12-month follow-up were rather similar, except that a high BMI was no longer associated with poor recovery. 'Confidence in recovery in general' ('no confidence' and 'do not know'), was now associated with poor recovery. The explained variance (Nagelkerke R^2) was 0.318 and 0.345 for 3 and 12 months, respectively. The percentage of patients correctly classified by the models was 73.3 and 75.2% for the 3 and 12-month follow-up, respectively. These results were rather robust: different RDQ values (≤ 3 and ≤ 5) only revealed that 'confidence in recovery in general' was no longer associated with the 12-month follow-up based on an RDQ score ≤ 5.

Table 3
Prognostic indicators of functional status (recovered: RDQ score ≤4)

Variable	3 Months				12 Months			
	Individual indicators		Multiple regression		Individual indicators		Multiple regression	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Functional status (RDQ >13)[a]	0.21	0.09; 0.49*	0.26	0.10; 0.67	0.25	0.11; 0.59*	0.49	0.19; 1.29
Gender (female) (*n*=45)	1.08	0.48; 2.43			0.56	0.25; 1.24*		
Age (≥43)[a]	0.59	0.26; 1.33*			0.80	0.37; 1.74		
Duration of complaints before surgery								
(1) ≥2–≤6 months (*n*=56)	1				1			
(2) ≥7–≤12 months (*n*=40)	1.16	0.49; 2.57			0.95	0.42; 2.14		
(3) >12 months (*n*=9)	0.48	0.09; 2.55			0.33	0.06; 1.74*		
Medication at baseline (yes) (*n*=38)	0.25	0.10; 0.64*	0.25	0.09; 0.73	0.16	0.06; 0.42*	0.15	0.05; 0.43
Days in hospital (>7 days) (*n*=28)	0.89	0.36; 2.21			1.14	0.47; 2.73		
Body mass index								
(1) <25 (*n*=43)	1		1		1			
(2) ≥25, <30 (*n*=47)	0.53	0.22; 1.27*	0.35	0.12; 0.99	0.71	0.31; 1.63		
(3) ≥30 (*n*=15)	0.83	0.25; 2.76	0.79	0.21; 2.94	0.70	0.21; 2.29		
Pain back (VAS >45)[a]	0.92	0.41; 2.03			0.80	0.37; 1.74		
Pan leg (VAS >43)[a]	0.79	0.36; 1.76			0.55	0.25; 1.20*		
Fear of movement (TSK >40) (*n*=77)	0.91	0.37; 2.25			0.63	0.26; 1.54		
Catastrophizing (PCS)								
(1) ≤11	1				1			
(2) >11–≤21	0.31	0.11; 0.84*			0.54	0.21; 1.38*		
(3) >21	0.35	0.13; 0.94*			0.36	0.14; 0.96*		
Negative affectivity (NEM)								
(1) ≤1	1				1			
(2) >1–≤4	0.46	0.17; 1.20*			0.75	0.30; 1.90*		
(3) >4	0.49	0.18; 1.32*			0.50	0.19; 1.33*		
Confidence in recovery								
(1) Great deal (*n*=43)	1				1		1	
(2) Moderate (*n*=42)	0.88	0.37; 2.11			0.50	0.21; 1.18*	0.80	0.29; 2.18
(3) No confidence (*n*=6)	0.25	0.03; 2.35			0.14	0.02; 1.34*	0.24	0.02; 2.63
(4) Do not know (*n*=14)	0.19	0.04; 3.29			0.19	0.05; 0.78*	0.22	0.05; 1.12
Patients' treatment expectancy (≥6) (*n*=80)	2.94	1.00; 8.64*	4.22	1.23; 14.42	2.44	0.92; 6.48*	3.03	0.98; 9.39
Nagelkerke R^2			0.318				0.345	
Correctly classified (0.500)			73.3%				75.2%	

Odds Ratios (OR) and 95% Confidence Interval (CI) for each indicator. All analyses adjusted for intervention. **P*-value ≤0.250 for individual indicator, included in multivariate logistic regression. Variables with more than two categories were included if only one category had a *P*-value ≤0.250. In multivariate logistic regression only variables with *P*-value ≤0.10. RDQ, Roland Disability Questionnaire; VAS, Visual Analogue Scale; TSK, Tampa Scale for Kinesiophobia; PCS, Pain Catastrophizing Scale; NEM, Negative Emotionality sub-scale.

[a] Cut-off value=median, for all other variables numbers in categories between brackets. For dichotomous variables the reference category is the contrast (female versus male), for all other variables the reference category is first category (OR=1).

3.3. Back pain

With respect to predictors of the severity of back pain, Table 4 shows that the intensity of back pain at baseline was associated with a poor outcome on back pain at the 3 and 12-month follow-up. Taking pain medication at baseline and negative affectivity were only associated with poor outcome at the 3-month follow-up, whereas poor functional status (RDQ >13) and leg pain at baseline were only associated with poor outcome at the 12-month follow-up.

The overall predictive power of the models was modest, with an explained variance (Nagelkerke R^2) of 0.297 and 0.143 for 3 and 12 months, respectively. The percentage of patients who were correctly classified by the models was 77.1 and 64.8% for the 3 and 12-month follow-up, respectively. The sensitivity analysis showed that using a cut-off value of ≤20 mm altered the results. Negative affectivity was no longer associated with back pain at the 3-month follow-up, but was replaced by treatment expectancy. At the 12-month follow-up 'RDQ>13' was replaced by pain catastrophizing. However, the overall predictive power of the model did not change substantially, and remained modest.

3.4. Leg pain

Table 5 shows that both leg pain at baseline and a poor functional status (RDQ>13) at baseline were associated with a poor outcome for leg pain at the 3 and 12-month follow-up. In addition, high treatment expectancy was

Table 4
Prognostic indicators of back-pain (recovered: VAS score $\leq$10 mm)

Variable	3 Months				12 Months			
	Individual indicators		Multiple regression		Individual indicators		Multiple regression	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Pain back (VAS >45)[a]	0.21	0.08; 0.52*	0.24	0.09; 0.63	0.69	0.32; 1.50	0.72	0.31; 1.65
Functional status (RDQ >13)[a]	0.42	0.17; 1.00*			0.32	0.14; 0.72*	0.35	0.15; 0.81
Gender (female) (n=45)	0.63	0.26; 1.51			0.67	0.30; 1.48		
Age ($\geq$43)[a]	0.38	0.16; 0.93*			0.57	0.26; 1.25*		
Duration of complaints before surgery								
(1) $\geq$2–$\leq$6 months (n=56)	1				1			
(2) $\geq$7–$\leq$12 months (n=40)	0.80	0.33; 1.95			1.90	0.84; 4.34*		
(3) >12 months (n=9)	0.58	0.11; 3.17			0.45	0.09; 2.37		
Medication at baseline (yes) (n=38)	0.31	0.11; 0.85*	0.28	0.10; 0.85	0.45	0.20; 1.04*		
Days in hospital (>7 days) (n=28)	0.75	0.28; 2.00			0.74	0.31; 1.81		
Body mass index								
(1) <25 (n=43)	1				1			
(2) $\geq$25, <30 (n=47)	1.37	0.55; 3.43			1.23	0.53; 5.67		
(3) $\geq$30 (n=15)	1.46	0.41; 5.24			1.73	0.53; 5.67		
Pain leg (VAS >43)[a]	0.79	0.34; 1.83			0.46	0.21; 1.02*	0.47	0.20; 1.08
Fear of movement (TSK >40) (n=77)	0.57	0.21; 1.58			0.77	0.32; 1.86		
Catastrophizing (PCS)								
(1) $\leq$11	1				1			
(2) >11–$\leq$21	0.61	0.27; 1.63			0.76	0.30; 1.95		
(3) >21	0.32	0.11; 0.95*			0.35	0.13; 0.94*		
Negative affectivity (NEM)								
(1) $\leq$1	1		1		1			
(2) >1–$\leq$4	0.46	0.17; 1.22*	0.55	0.19; 1.61	0.93	0.37; 2.34		
(3) >4	0.16	0.05; 0.56*	0.21	0.06; 0.78	0.57	0.21; 1.50*		
Confidence in recovery								
(1) Great deal (n=43)	1				1			
(2) Moderate (n=42)	0.49	0.20; 1.23*			0.55	0.23; 0.30*		
(3) No confidence (n=6)	0.28	0.03; 2.59*			0.43	0.07; 2.64		
(4) Do not know (n=14)	0.11	0.01; 0.90*			0.45	0.13; 1.60*		
Patients' treatment expectancy ($\geq$6) (n=80)	3.97	1.09; 14.46*			1.52	0.60; 3.84		
Nagelkerke R^2			0.297				0.143	
Correctly classified (0.500)			77.1%				64.8%	

Odds Ratios (OR) and 95% Confidence Interval (CI) for each indicator. All analyses adjusted for intervention. *P-value $\leq$0.250 for individual indicator, included in multivariate logistic regression. Variables with more than two categories were included if only one category had a P-value $\leq$0.250. In multivariate logistic regression only variables with P-value $\leq$0.10. RDQ, Roland Disability Questionnaire; VAS, Visual Analogue Scale; TSK, Tampa Scale for Kinesiophobia; PCS, Pain Catastrophizing Scale, NEM, Negative Emotionality sub-scale.

[a] Cut-off value=median, for all other variables numbers in categories between brackets. For dichotomous variables the reference category is the contrast (female versus male), for all other variables the reference category is first category (OR=1).

associated with a positive outcome for leg pain at the 3-month follow-up.

The overall predictive power of the models was modest: Nagelkerke R^2 of 0.226 and 0.197 for 3 and 12 months, respectively. The percentage of patients who were correctly classified was 68.6 and 69.5% for the 3 and 12-month follow-up, respectively. As with back pain, a cut-off value of $\leq$20 mm altered the results slightly. At the 3-months follow-up only leg pain at baseline was retained in the multivariate model. At the 12-month follow-up treatment expectancy was no longer retained in the multivariate model. Also, as with back pain, the overall predictive power did not change substantially, and remained modest.

4. Discussion

The aim of our study was to identify indicators of the short and long-term outcome of residual complaints following lumbar disc surgery that can easily be measured in primary care. Not surprisingly, baseline scores for the outcome measures are strong predictors of the respective outcomes of the 3 and 12-month follow-up measurements. Poor functional status (RDQ>13) at baseline was associated with a poor outcome on functional status, and the baseline scores for leg pain and back pain predicted the outcomes on leg pain and back pain, respectively. Taking pain medication was associated with a poor outcome on both follow-up measurements of perceived recovery

Table 5
Prognostic indicators of leg pain (recovered: VAS score ≤10 mm)

Variable	3 Months				12 Months			
	Individual indicators		Multiple regression		Individual indicators		Multiple regression	
	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Pan leg (VAS>43)[a]	0.25	0.11; 0.58*	0.24	0.10; 0.58	0.38	0.17; 0.85*	0.38	0.16; 0.75
Pain back (VAS>45)[a]	1.04	0.48; 2.28			1.44	0.66; 3.15		
Functional status (RDQ>13)[a]	0.40	0.18; 0.91*	0.45	0.18; 1.08	0.44	0.19; 0.98*	0.31	0.13; 0.75
Gender (female) (n=45)	0.80	0.36; 1.77			1.04	0.47; 2.29		
Age (≥43)[a]	0.73	0.33; 1.61			1.05	0.48; 2.28		
Duration of complaints before surgery								
(1) ≥2–≤6 months (*n*=56)	1				1			
(2) ≥7–≤12 months (*n*=40)	1.27	0.56; 2.88			1.01	0.45; 2.29		
(3) >12 months (*n*=9)	0.78	0.18; 3.43			0.83	0.20; 3.47		
Medication at baseline (yes) (*n*=38)	0.45	0.19; 1.04*			0.64	0.28; 1.44		
Days in hospital (>7 days) (*n*=28)	0.59	0.24; 1.48			0.77	0.32; 1.87		
Body mass index								
(1) <25 (*n*=43)	1				1			
(2) ≥25, <30 (*n*=47)	1.36	0.59; 3.17			0.82	0.35; 1.89		
(3) ≥30 (*n*=15)	1.12	0.34; 3.75			0.56	0.17; 1.87		
Fear of movement (TSK>40) (*n*=77)	1.11	0.46; 2.67			1.00	0.42; 2.40		
Catastrophizing (PCS)								
(1) ≤11	1				1			
(2) >11–≤21	1.31	0.51; 3.35			1.11	0.43; 2.86		
(3) >21	0.80	0.31; 2.09			0.51	0.20; 1.34*		
Negative affectivity (NEM)								
(1) ≤1	1				1			
(2) >1–≤4	0.67	0.26; 1.70			1.02	0.40; 2.58		
(3) >4	0.55	0.21; 1.47*			0.78	0.30; 2.03		
Confidence in recovery								
(1) Great deal (*n*=43)	1				1			
(2) Moderate (*n*=42)	0.59	0.25; 1.41*			0.51	0.21; 1.22*		
(3) No confidence (*n*=6)	0.91	0.02; 1.77*			0.65	0.12; 3.70		
(4) Do not know (*n*=14)	0.38	0.10; 1.40*			0.31	0.09; 1.12		
Patients' treatment expectancy (≥6) (*n*=80)	2.72	0.98; 7.54*	3.00	1.00; 9.00	1.32	0.53; 3.29		
Nagelkerke R^2			0.226				0.197	
Predicted (0.500)			68.6%				69.5%	

Odds Ratios (OR) and 95% Confidence Interval (CI) for each indicator. All analyses adjusted for intervention. **P*-value ≤0.250 for individual indicator, included in multivariate logistic regression. Variables with more than two categories were included if only one category had a *P*-value ≤0.250. In multivariate logistic regression only variables with *P*-value ≤0.10. RDQ, Roland Disability Questionnaire; VAS, Visual Analogue Scale; TSK, Tampa Scale for Kinesiophobia; PCS, Pain Catastrophizing Scale; NEM, Negative Emotionality sub-scale.

[a] Cut-off value=median, for all other variables numbers in categories between brackets. For dichotomous variables the reference category is the contrast (female versus male), for all other variables the reference category is first category (OR=1).

and functional status and with a poor outcome for back pain at the 3-month follow-up. Taking medication at baseline seems to be an indicator of the severity of complaints, because it was highly correlated with leg pain and a poor functional status at baseline. Patients taking pain medication had a statistically significant higher RDQ score ($P<0.000$), and scored statistically significantly higher ($P=0.025$) on leg pain at baseline. The multivariate models revealed that this indicator was highly significantly associated with perceived recovery and functional status, also when adjusted for leg pain and decreased functional status.

For functional status and perceived recovery at 3 and 12 months, and for leg pain at 3 months, a favorable prognosis was found for patients with high treatment expectancy. It has been reported that treatment expectancy is a significant predictor of the outcome of treatment for anxiety disorders or depression (Kirsch and Sapirstein, 1999), in the magnitude of the effect of placebo analgesia (Price et al., 1999) and the outcome of conservative treatment for low back pain (Kalauokalani et al., 2001). Moreover, treatment expectancy predicts the outcome of surgical treatment for sciatica (Lutz et al., 1999). The results of our study seem to support these findings, but it should be noted that the overall predictive power of the models was modest, at best. There might be some overlap between treatment expectancy and the patient's 'confidence in recovery in general' but the results show that the predictive value of both concepts is rather different. We think that the main reason for this is that treatment expectancy specifically focuses on the treatment, and is based on actual perception, as opposed to 'confidence in recovery in general'. There is still no unambiguous explanation for the effects treatment expectancy, and

various mechanisms could apply (Lurie, 2001). It has been suggested that treatment expectancy is partly responsible for the 'non-specific' effects of treatment (Deyo et al., 1990; Luparello et al., 1970; Price et al., 1999). However, from a methodological point of view there might be some concerns with regard to the assessment of expectancy. As stated, rating of treatment expectancy was based on the actual perception of the treatment, because it was rated after the first two treatment sessions. Therefore in our study treatment expectancy includes two components: expectancy with regard to the treatment effect and the patient–therapist interaction. However, no psychometric properties have yet been established for any of the methods used to assess treatment expectancy. Recently Devilly (Devilly and Borkovec, 2000) assessed the psychometric properties of the credibility/expectancy questionnaire (CEQ), and although they seem to be promising, the authors still stress the need for caution using the scale.

Contrary to what we had anticipated, fear of movement and pain-catastrophizing were not associated with any of the dimensions of recovery. None of the univariate models demonstrated a significance level of <0.250. Therefore, these indicators were not included in the multivariate models. A possible explanation could be that the TSK and the PCS are not specifically focused on a post-surgery population. This population might have fear of some very specific movements, such as turning over in bed or lifting heavy objects from the floor, whereas the TSK items only relate to general activities. In other words, further research is necessary to validate these measures in a post-surgery population. In order to assess the robustness of the results we re-analyzed our data using a cut-off value of >42 (as opposed to 40). However, the different cut-off value made no substantial change in the results.

It could be argued that measuring some variables pre-operatively is important, but in this study we focused on prognostic variables for residual complaints following lumbar disc surgery. In our opinion, it is more appropriate to measure post-operatively variables, such as fear of movement and pain catastrophzing that are likely to be altered by the surgery. Moreover, treatment expectancy was rated after two treatment sessions because we think that it is important that patients can rate their expectancy based on their actual perception of the treatment, and not simply on their assumptions. Initially, we also planned to include some variables that were measured before the surgery (e.g. muscle tendon reflexes, pain intensity) but due to practical problems this was not possible. Measuring the Laseque was not informative, because a positive pre-surgery Laseque was one of the inclusion criteria for the surgery. To account for the severity of the pre-surgery complaints, the duration of complaints before surgery was included as a potential indicator because this was measured by patient self-reports. However, it was not found to be associated with any of the outcomes in this study. The type of surgery was not included in the analyses, because the majority of the patients underwent a standard discectomy. It is therefore not know whether our results also apply to other types of surgery for herniated lumbar discs, but as the standard discectomy is the most frequently applied technique our study population represents the majority of post lumbar disc surgery patients.

Our sample of patients was studied prospectively, so this can be considered as a cohort study in which the interventions are regarded as prognostic indicators. It appeared that there were no differences between the two intervention groups in either treatment expectations or outcomes. Although there were no differences between the two groups, the range within the outcome measures (as indicated by relatively large SDs) was broad. However, the range within treatment expectancy was narrow which might have affected the predictive power of this measurement. If there had been a broader range in treatment expectancy we expect that the predictive power would have been stronger. However, there are no indications that this narrow range in scores in treatment expectancy is a result of the method of patient recruitment. The inclusion and exclusion criteria for our trial (and thus for referring patients to primary care treatment) reflect the daily clinical practice of the neurosurgeon during their 6 weeks consult. Therefore we do not think that our sample was selected on the basis of prognosis. Certainly not all patients with residual complaints were referred by the neurosurgeon. The majority of patients who were potentially eligible, but did not participate in the study were excluded due to practical reason, as described in the first paragraph of Section 3. Therefore we think that a broad spectrum of patients with residual complaints was included.

We used dichotomized outcome measures because we focused on indicators for absolute recovery, and not just improvement (or change) in pain or functional status. However, the use of cut-off values is somewhat arbitrary, so we performed sensitivity analyses in order to assess the robustness of the results. For functional status the results proved to be robust: using different RDQ scores (≤ 3 and ≤ 5) as cut-off values at follow-up yielded similar findings. However, for leg pain and back pain, the results were less stable. Using different cut-off values yielded different multivariate models, but the overall predictive power of the different models remained modest, at best. Apparently, the clinical course to recovery of residual leg pain and residual back pain is not strongly influenced by these indicators.

Finally, the results of our study have some relevance for application in primary care. First of all, a few indicators could help health care providers to identify patients with a higher risk, but more high quality studies are needed to investigate a broad spectrum of potential indicators, because the overall predictive value of the models was low. Secondly, the role of treatment expectancy is interesting. Our results lent some support to the assumption that, in addition to being evidence based, interventions that are offered to patients should preferably also meet the patient's

expectations, whenever possible. Moreover, health care providers could play an important role in positively altering a patient's treatment expectancy. It has already been demonstrated that treatment expectancy can be greatly influenced by the amount and length of the information provided, and also by the kind of language used to communicate the information (Horvath, 1990; Kazdin and Krouse, 1983).

References

Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. J Behav Ther Exp Psychiatry 2000;31:73–86.

Deyo RA, Walsh NE, Martin DC, Schoenfeld LS, Ramamurthy S. A controlled trial of transcutaneous electrical nerve stimulation (TENS) and exercise for chronic low back pain. N Engl J Med 1990;322: 1627–34.

Dvorak J, Valach L, Fuhrimann P, Heim E. The outcome of surgery for lumbar disc herniation. II. A 4–17 years' follow-up with emphasis on psychosocial aspects. Spine 1988;13:1423–7.

Fritz JM, George SZ, Delitto A. The role of fear-avoidance beliefs in acute low back pain: relationships with current and future disability and work status. Pain 2001;94:7–15.

Horvath P. Treatment expectancy as a function of the amount of information presented in therapeutic rationales. J Clin Psychol 1990; 46:636–42.

Hurme M, Alaranta H. Factors predicting the result of surgery for lumbar intervertebral disc herniation. Spine 1987;12:933–8.

Kalauokalani D, Cherkin DC, Sherman KJ, Koepsell TD, Deyo RA. Lessons from a trial of acupuncture and massage for low back pain: patient expectations and treatment effects. Spine 2001;26:1418–24.

Kazdin AE, Krouse R. The impact of variations in treatment rationales on expectancies for therapeutic change. Behav Ther 1983;14:657–71.

Kirsch I, Sapirstein G. Listening to prozac but hearing placebo: a meta-analysis of antidepressant medication. In: Kirsch I, editor. How expectancies shape experience. Washington, DC: American Psycholoigical Association; 1999. p. 033–320.

Kjellby-Wendt G, Styf JR, Carlsson SG. The predictive value of psychometric analysis in patients treated by extirpation of lumbar intervertebral disc herniation. J Spinal Disord 1999;12:375–9.

Kori SH, Miller RP, Todd DD. Kinisophobia: a new view of chronic pain behavior. Pain Manage 1990;35–43.

Korres DS, Loupassis G, Stamos K. Results of lumbar discectomy: a study using 15 different evaluation methods. Eur Spine J 1992;1:20–4.

Luparello TJ, Leist N, Lourie CH, Sweet P. The interaction of psychologic stimuli and pharmacologic agents on airway reactivity in asthmatic subjects. Psychosom Med 1970;32:509–13.

Lurie JD. Point of view. Spine 2001;26:1424.

Lutz GK, Butzlaff ME, Atlas SJ, Keller RB, Singer DE, Deyo RA. The relation between expectations and outcomes in surgery for sciatica. J Gen Intern Med 1999;14:740–4.

Manniche C, Asmussen KH, Vinterberg H, Rose Hansen EB, Kramhoft J, Jordan A. Analysis of preoperative prognostic factors in first-time surgery for lumbar disc herniation, including Finneson's and modified Spengler's score systems. Dan Med Bull 1994a;41:110–5.

Manniche C, Asmussen KH, Vinterberg H, Rose Hansen EB, Kramhoft J, Jordan A. Back pain, sciatica and disability following first-time conventional haemilaminectomy for lumbar disc herniation. Use of 'low back pain rating scale' as a postal questionnaire. Dan Med Bull 1994b;41:103–6.

Ostelo RW, de Vet HC, Vlaeyen JW, Kerckhoffs MR, Berfelo WM, Wolters PM, Van Den Brandt PA. Behavioral graded activity following first-time lumbar disc surgery: 1-year results of a randomized clinical trial. Spine 2003a;28:1757–65.

Ostelo RW, de Vet HC, Waddell G, Kerckhoffs MR, Leffers P, van Tulder MW. Rehabilitation after lumbar disc surgery. Spine 2003b;28: 209–18.

Pappas CTE, Harrington T, Sonntag VKH. Outcome analysis in 654 surgically treated lumbar disc herniations. Neurosurgery 1992;30: 862–6.

Picavet HS, Vlaeyen JW, Schouten JS. Pain catastrophizing and kinesiophobia: predictors of chronic low back pain. Am J Epidemiol 2002;156:1028–34.

Price DD, Milling LS, Kirsch I, Duff A, Montgomery GH, Nicholls SS. An analysis of factors that contribute to the magnitude of placebo analgesia in an experimental paradigm. Pain 1999;83:147–56.

Roland M, Morris R. 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 1983;8:141–4.

Stegen K, Neujens A, Crombez G, Hermans D, Van de Woestijne KP, Van den Bergh O. Negative affect, respiratory reactivity, and somatic complaints in a CO2 enriched air inhalation paradigm. Biol Psychol 1998;49:109–22.

Sullivan MJL, Bishop SR, Pivik J. The pain catastrophizing scale: development and validation. Psychol Assess 1995;7:524–32.

Tellegen A. Brief manual for the multidimensional personality questionnaire. Unpublished manual. University of Minnesota; 1982.

Vlaeyen JW, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain 2000;85:317–32.

Vlaeyen JW, Teeken Gruben RA, van Eek H, Heuts PH. N.J., Goossens, M.E., Rutten van Molken, M.P., Pelt, Cognitive-educational treatment of fibromyalgia: a randomized clinical trial. I. Clinical effects. J Rheumatol 1996;23:1237–45.

Weber H. Lumbar disc herniation. A controlled, prospective study with ten years of observation. Spine 1983;8:131–40.

Yorimitsu E, Chiba K, Toyama Y, Hirabayashi K. Long-term outcomes of standard discectomy for lumbar disc herniation: a follow-up study of more than 10 years. Spine 2001;26:652–7.