Eur Spine J (2012) 21:681–690
DOI 10.1007/s00586-011-2145-3

ORIGINAL ARTICLE

Predictors of outcome after surgery with disc prosthesis and rehabilitation in patients with chronic low back pain and degenerative disc: 2-year follow-up

Christian Hellum · Lars Gunnar Johnsen · Øyvind Gjertsen · Linda Berg · Gesche Neckelmann · Oliver Grundnes · Ivar Rossvoll · Jan Sture Skouen · Jens Ivar Brox · Kjersti Storheim · The Norwegian Spine Study Group

Received: 21 July 2011 / Revised: 15 December 2011 / Accepted: 31 December 2011 / Published online: 13 January 2012

Abstract
Purpose A prospective study to evaluate whether certain baseline characteristics can predict outcome in patients treated with disc prosthesis or multidisciplinary rehabilitation.
Methods Secondary analysis of 154 patients with chronic low back pain (LBP) for at least 1 year and degenerative discs originally recruited for a randomized trial. Outcome measures were Oswestry Disability Index (ODI) dichotomized to $<$ or $\geq$15 points improvement and whether subjects were working at 2-year follow-up. A multiple logistic regression analysis was used.
Results In patients treated with disc prosthesis, long duration of LBP and high Fear-Avoidance Beliefs for work (FABQ-W) predicted worse ODI outcome [odds ratio (OR) = 1.9, 95% confidence interval (CI) 1.2–3.2 and OR = 1.7, CI 1.2–2.4 for every 5 years or 5 points]. Modic type I or II predicted better ODI outcome (OR = 5.3, CI 1.1–25.3). In patients treated with rehabilitation, a high ODI, low emotional distress (HSCL-25), and no daily narcotics predicted better outcome for ODI (OR = 2.5, CI 1.4–4.5 for every 5 ODI points, OR = 2.1, CI 1.1–5.1 for every 0.5 HSCL points and OR = 23.6, CI 2.1–266.8 for no daily narcotics). Low FABQ-W and working at baseline predicted working at 2-year follow-up after both treatments (OR = 1.3, CI 1.0–1.5 for every 5 points and OR = 4.1, CI 1.2–13.2, respectively).
Conclusions Shorter duration of LBP, Modic type I or II changes and low FABQ-W were the best predictors of success after treatment with disc prosthesis, while high ODI, low distress and not using narcotics daily predicted

Additional investigators of the Norwegian Spine Study Group who participated in the study are listed in the "Appendix."

C. Hellum (✉) · J. I. Brox
Department of Orthopaedics, Oslo University Hospital, University of Oslo, Kirkevn 166, 0407 Oslo, Norway
e-mail: christian.hellum@medisin.uio.no

L. G. Johnsen · I. Rossvoll
National Centre for Diseases of the Spine, Trondheim University Hospital, Trondheim, Norway

L. G. Johnsen · I. Rossvoll
Orthopaedic Department, Trondheim University Hospital, Trondheim, Norway

Ø. Gjertsen
Department of Radiology, Oslo University Hospital, University of Oslo, Oslo, Norway

L. Berg · G. Neckelmann
Department of Radiology, Haukeland University Hospital, Bergen, Norway

L. Berg
Section for Radiology, Department of Surgical Sciences, University of Bergen, Bergen, Norway

L. Berg
Curato Røntgen Bodø, Bodø, Norway

O. Grundnes · K. Storheim
Hjelp24, Nimi, Oslo, Norway

J. S. Skouen
The Outpatient Spine Clinic, Haukeland University Hospital, University of Bergen, Bergen, Norway

K. Storheim
NAR, Department of Orthopaedics, Oslo University Hospital, University of Oslo, Kirkevn 166, 0407 Oslo, Norway

K. Storheim
Communication-and research unit for musculoskeletal disorders, FORMI, Oslo University Hospital, Oslo, Norway

better outcome of rehabilitation. Low FABQ-W and working predicted working at follow-up.

Keywords LBP · Degenerative disc · Disc replacement · Multidisciplinary rehabilitation

Introduction

Fusion has been the traditional surgical treatment option for patients with low back pain (LBP) and radiographic signs of degeneration. Since the middle of the 1980s disc prostheses have been introduced as a treatment option [1–5]. However, in the same period, non-surgical interventions have also been shown to be effective in treating patients with chronic LBP, especially multidisciplinary rehabilitation interventions focusing on physical exercise and/or cognitive behavioral principles [6–9].

There is limited knowledge about valid predictors of good or poor outcomes in patients treated with multidisciplinary rehabilitation, fusion and especially disc prosthesis [10]. Improving the ability to predict treatment outcome may allow for better allocation of resources. Several authors have evaluated predictors of treatment outcome, but the few prospective studies and the many retrospective studies conducted are heterogeneous. Populations, outcome measures and statistical methods differ greatly from study to study, resulting in a multitude of possible predictors. In summary, personality traits and labor-related factors seem, to some degree, to be consistent predictors [7, 11].

The aim of the present study was to evaluate whether any physical, socio-demographic, psychological/pain and radiological characteristics at baseline could predict improvement in Oswestry Disability Index (ODI) and return-to work at 2-year follow-up in patients treated with disc prostheses or with multidisciplinary rehabilitation.

Patients and methods

Data were extracted from a randomized study conducted at five Norwegian university hospitals [12]. The study was evaluated and approved by the Regional Committees for Medical Research Ethics, conducted in accordance with the Helsinki Declaration and the ICH-GCP guidelines and registered at http://www.clinicaltrial.gov under the identifier (NCT 00394732).

Design

The present study was an analysis of patients who completed a rehabilitation program or underwent disc prosthesis surgery and either attended the 2-year follow-up or filled in a postal questionnaire. Patients crossing over from rehabilitation to surgery ($n = 5$) were analyzed in the surgical group and patients not completing the rehabilitation program were taken out of the analysis according to as-treated principles.

Participants

Inclusion criteria were age 25–55 years, LBP as the main symptom for at least 1 year, ODI score $\geq$30%, and degenerative changes in the intervertebral disc at one or both of the lower lumbar levels (L4/L5 and/or L5/S1). For further details see Hellum et al. [12].

Study interventions

Rehabilitation was based on the multidisciplinary treatment model described by Brox et al. [6], and consisted of a cognitive approach and supervised physical exercise.

The surgical intervention consisted of replacement of the degenerative intervertebral lumbar disc with an artificial lumbar disc (ProDisc II, Synthes Spine, West Chester, PA, USA) [12].

Dependent variables

The primary outcome measure was the change in functional capacity from baseline to 2-year follow-up, measured by ODI (version 2.0) [13]. Change in ODI was dichotomized to $<$ or $\geq$15 points improvement. The cut off value was chosen based on data from the Food and Drug Administration considering an individual change in ODI of more than 15 ODI points as the minimal clinically important change [12].

A second analysis was performed regarding work status. Patients who were working part or full time at 2-year follow-up were categorized as working, likewise students and homemakers. Data was collected from the patients and from the National Insurance of employees (NAV).

Baseline variables tested for predictive value (independent variables)

Potential predictors were registered at baseline and grouped into physical, socio-demographic, psychological variables and pain, and radiological variables (Table 1).

Physical variables. Daily consumption of narcotics, prior surgery, level(s) operated on, the presence of comorbidity, ODI, back pain duration and body mass index (BMI).

Sociodemographic variables. Socioeconomic status was classified according to the Norwegian Standard

Table 1 Step 1. Baseline variables associated with or without 15 ODI points improvement at follow-up in the surgical and nonsurgical groups separately

	Surgery ($n = 88$)			Rehabilitation ($n = 66$)		
	N (yes/no)	% patients improved ≥15 ODI points categorized by yes/no	*P* value	*N* (yes/no)	% patients improved ≥15 ODI points categorized by yes/no	*P* value
Physical variables						
Daily consumption of narcotics	27/61	72/70	0.87¶	9/57	22/49	0.17¶
Prior surgery	27/61	74/67	0.50†	18/48	39/48	0.51†
Affected level						
L4/L5	19/69	83/68	0.29†	13/53	46/45	0.98†
L5/S1	45/43	64/79		37/29	46/44	
L4/L5 and L5/S1	24/64	75/70		16/50	43/46	
Comorbidity	20/68	65/74	0.46†	18/48	50/44	0.65†
ODI	42 ± 9.2[a]		0.59	42 ± 8.1[a]		0.03*
Duration of back pain (years)	6.6 ± 6.1[a]		0.14‡	7.1 ± 6.4[a]		0.89‡
Body mass index	25.6 ± 3.3[a]		0.08*	25.6 ± 3.4[a]		0.19*
Sociodemographic variables						
Socioeconomic status						
Manual worker[b]	35/49	60/80	0.05†	23/37	44/49	0.70†
Educational level						
Primary school (9 years)	20/68	60/75	0.25†	12/54	58/43	0.54†
High school (12 years)	47/41	79/63		44/22	41/55	
University/college[c]	21/67	67/73		10/56	50/45	
Working[d]	22/66	82/68	0.22†	12/54	75/39	0.023†
Duration out of work (months)						
<6 months	23/61	87/67	0.17†	18/38	67/37	0.08†
6 months–1 year	18/66	72/73		18/38	44/47	
>1 year	43/41	65/80		20/36	30/56	
Sex (female)	43/45	79/64	0.13†	38/28	47/43	0.72†
Current smoker	42/45	71/71	0.97†	29/37	35/54	0.11†
Age	40.9 ± 7.1[a]		0.27*	41.5 ± 6.9[a]		0.74*
Psychological variables and pain						
HSCL	1.8 ± 0.5[a]		0.43*	1.9 ± 0.5[a]		0.06*
FABQ-work	26.5 ± 10.6[a]		0.01*	27.8 ± 9.8[a]		0.61*
FABQ-physical	13.1 ± 5.7[a]		0.89*	12.0 ± 5.7[a]		0.96*
MCS (SF-36)	46.5 ± 13.1[a]		0.13*	45.2 ± 13.4[a]		0.14*
Back pain (VAS)	69.2 ± 15.2[a]		0.75*	73.2 ± 13.0[a]		0.33*
Pain drawing (below belt)[e]	67/14	73/71	1.0†	*44/15*	41/53	0.40†
Radiological variables						
Modic I and II						
Not present	13/74	39/77	0.04†	9/59	25/49	0.37†
Modic primary type I	26/61	81/67		22/44	41/49	
Modic primary type II	33/54	76/70		22/44	59/40	
Modic I and II	15/72	73/71		13/53	46/46	
Modic CC[f]	31/56	81/66	0.15†	25/40	52/43	0.46†
Disc height reduction[g]	63/24	71/71	0.96†	45/20	49/40	0.51†
Nucleus pulposus grade 3/4	78/9	69/89	0.22†	55/10	47/40	0.67†
Facet joint arthropathy grade 2/3	10/77	70/71	0.93†	4/61	75/44	0.23†

Table 1 continued

	Surgery ($n = 88$)			Rehabilitation ($n = 66$)		
	N (yes/no)	% patients improved ≥15 ODI points categorized by yes/no	*P* value	*N* (yes/no)	% patients improved ≥15 ODI points categorized by yes/no	*P* value
Posterior HIZ	41/46	68/74	0.56[†]	23/42	39/50	0.40[†]

Continuous and categorical variables

The Oswestry Disability Index (ODI) ranges from 0 to 100, with lower scores indicating less severe symptoms

Back pain was calculated using a horizontal scale ranging from 0 (no pain) to 100 (worst pain imaginable) with word anchors at the beginning and end

SF-36 scores range from 0 to 100, with higher scores corresponding to better health status

Waddell's FABQ scale ranges from 0 to 24 (physical) and from 0 to 42 (work), with lower scores indicating less severe symptoms

P values; Indicating if the baseline variable was associated with 15 ODI points improvement or not at follow up

HSCL-25 = Hopkins symptom checklist; HSCL-25 reflects emotional distress and scores range from 1 to 4, with lower scores indicating less severe symptoms

FABQ Fear of Avoidance Belief Questionnaire

* Independent two-sided *t* test; [†]Pearson's χ^2; [‡]Mann–Whitney *U* test; [¶]Fischer's exact test

[a] Values are represented as mean ± SD

[b] Classified according to socioeconomic status from Statistics Norway consisting of six groups: manual low, manual high, routine nonmanual low, routine nonmanual high, professional low or professional high. Because there were few patients in each group, unskilled and skilled workers were merged into the same group and nonmanual in one group. Consequently, two groups were analyzed, manual and nonmanual [14]

[c] Because there were fewer than five subjects in the category of education at university level, they were merged with 13–15 years of education (college)

[d] Working versus not working; Including part-time work as working

[e] Uden et al. [19]

[f] More than 50% of the craniocaudal diameter

[g] More or less than 40% height reduction (more = yes)

Classification of Socioeconomic status [14] and education according to the International Standard of Classification of Education (IECED) [15].Working status was categorized into working/not working, duration of sick leave categorized as <6 months, 6–12 months and >12 months, sex, smokers and age.

Psychological variables and pain scales. Emotional distress (Hopkins symptom checklist, HSCL-25), [16] Fear-Avoidance Beliefs Questionnaire (FABQ) [17], the Mental Component Scale (MCS) of SF-36 [18], LBP (VAS) and pain drawing [19].

Radiological variables. The classification/evaluation is presented in Table 2 [20–25]. Evaluation of the MRI examinations was performed independently by three experienced radiologists. The radiologists were blinded to clinical data. The outcome was decided by simple majority, by mean value or by a fourth radiologist when majority or mean was unsuitable (Modic type). In the univariate analysis of Modic changes, there was no difference between Modic primary type I and primary type II changes. These changes were merged to one category in the further multivariate analysis. The intra and inter observer reliability of the MRI evaluation will be published soon by our group.

We also evaluated patients with more or less than 55 ODI points at baseline following a recommendation by Prof. Jeremy Fairbank (personal communication).

Statistical analysis

All data were analyzed using SPSS (version 16, SPSS Inc., Chicago, IL, USA). Patients treated with surgery or rehabilitation were analyzed in separate cohorts for ODI as outcome, but merged for return to work as outcome.

We used a multiple logistic regression analysis in accordance with the purposeful selection model [26]. The main steps were as follows.

Step 1 *Univariate analysis*. Physical, socio-demographic, psychological/pain and radiological characteristics were analyzed in separate groups at this stage and merged in step 2 (Tables 1, 3).

For categorical data, Pearson's χ^2 and Fischer's exact test were used. Mann–Whitney independent sample tests were performed for continuous nonparametric variables and independent *t* test for continuous parametric variables. The *P* value was set to 0.20 according to the recommendation of Hosmer et al. [26].

Table 2 MRI evaluation at baseline

Radiological finding
Modic primary type I or II[a]
0: not present
1: Modic primary type I present
2: Modic primary type II present
3: Modic I and II present
Modic primary or secondary type I and/or II craniocaudal extension >50% of vertebral body height on sagittal images[b]
0: not present
1: present
Disc height reduction measured on midsagittal image[c]
0: less than 40% reduction
1: equal to or more than 40% reduction
Facet joint arthropathy[d]
0: No or slight arthropathy
1: Moderate or severe arthropathy
Posterior high-intensity zone (HIZ)[e]
0: not present
1: present
Nucleus pulposus signal grade 3/4[f]
0: not present
1: present

According to clinical practice, most MRI examinations consisted of sagittal T-2 and T-1 sequences, and axial T2, T1 or PD (proton density)

[a] Modic et al. [20]

[b] Jensen et al. [24]

[c] Masharawi et al. [25]. Measured in the midsagittal MRI compared with a normal disc above

[d] Fujiwara et al. [21]. Merging no arthrosis with slight arthrosis (grade 0 and 1) and moderate with severe arthrosis (grade 2 and 3)

[e] Aprill and Bogduk [22]

[f] Luoma et al. [23]

Step 2 *Multivariate analysis*. We used Wald statistics to exclude variables that did not seem important [26]. The significance level was set at $P < 0.05$. Thereafter, we reduced the model using the likelihood-ratio test to determine which variables should be included in the model. Further, we evaluated possibly significant or important confounders and added, one at a time, variables excluded from the initial multivariable model. Subsequently, the correct parametric form for continuous variables was identified and checked for plausible interactions. "First categorical" was chosen as the reference cell for variables with more than two categories.

Step 3. The final model in the main analysis representing the variables considered to be of predictive value (Tables 4, 5).

Step 4. The final model was tested for goodness-of-fit (Hosmer–Lemeshow test).

Age and sex did not have a confounding effect and were therefore not adjusted for through the analysis. Odds ratios (ORs) for continuous variables are reported in cluster; for every 5-point change in FABQ-W and ODI, for every 0.5 points of HSCL and for every 5-year duration of LBP.

Results

Of the 173 patients included in the original randomized study, in this secondary analysis, 154 were included whereof 66 patients completed the whole rehabilitation program and 88 patients underwent surgery with disc prosthesis. A flow chart presented previously illustrates the different reasons for not including the remaining 19 patients in the present analysis [12]. All models had acceptable goodness-of-fit tests (*P* value ranged from 0.38 to 0.41). None of the potential predictors had a Spearman's correlation coefficient >0.7.

Outcome

Surgery. Long duration of back pain and high FABQ-W predicted having an ODI change <15 points in the final model (OR = 1.9, confidence interval (CI) 1.2–3.2 and OR = 1.7, CI 1.2–2.4) (Table 4). This suggests that the odds of having an ODI change <15 was doubled for every 5-year duration of back pain and 1.7 for every 5 points of FABQ-W. The association between duration of back pain, FABQ-W and outcome was linear in logit. Modic primary or secondary type I and/or type II predicted better outcome (OR = 5.3, CI 1.1–25.3) (Table 4).

Rehabilitation. Not using narcotics daily (OR = 23.6 CI 2.1–266.8), high ODI at baseline (OR = 2.5, CI 1.4-2.5 for every 5-point increase in ODI) and low HSCL (OR = 2.4, CI 1.1–5.1 for every 0.5-point reduction), predicted having a change ≥15 points in ODI in the final model (Table 4).

Work (merged cohorts). Working at baseline predicted being at work at follow-up (OR = 4.1, CI 1.2–13.2). High FABQ-W was predictive for not being at work at 2-year follow-up with an odds ratio of 1.3 for every 5 points of FABQ-W (OR = 1.3, CI 1.0–1.5) (Table 5).

ODI (merged cohorts). Figure 1 illustrates the difference in outcome between surgery and rehabilitation in patients with high and lower levels of ODI (ODI <55 points and ODI ≥55 points (Fig. 1a, b). For patients with high levels of ODI at baseline ($n = 21$) we saw no significant difference in outcome between treatment groups.

Discussion

The choice of baseline variables was based on careful selection of plausible clinical predictors of outcome and

Table 3 Step 1. Baseline variables associated with working or not at follow-up in the surgical and nonsurgical groups

	Surgery and rehabilitation ($n = 154$)		
	N (yes/no)	% patients working categorized by yes/no	*P* value
Physical variables			
Daily consumption of narcotics	34/120	56/60	0.67†
Prior surgery	42/112	55/61	0.50†
Level			
L4/L5	32/122	69/57	0.41†
L5/S1	81/73	58/60	
L4/L5 and L5/S1	41/113	54/61	
Comorbidity	31/123	48/62	0.18†
ODI			0.02*
Duration of back pain (months)			0.97‡
Body mass index			0.25*
Sociodemographic variables			
Socioeconomic status			
Manual worker[a]	58/86	51/63	0.15†
Educational level (%)			
Primary school (9 years)	32/122	47/62	0.20†
High school (12 years)	93/61	60/57	
University/college[b]	29/125	69/57	
Working[c]	35/119	87/50	<0.001†
Duration out of work			
<6 months	43/97	79/54	0.009†
6 months–1 year	38/102	61/62	
>1 year	59/81	49/70	
Sex (female)	80/74	59/60	0.93†
Current smoker (%)	71/82	55/62	0.36†
Age			0.53*
Psychological variables and pain			
HSCL			0.52*
FABQ–work			0.001*
FABQ–physical			0.88*
MCS (SF-36)			0.25*
Back pain (VAS)			0.50*
Pain drawing (below belt)[d]	111/30	59/67	0.42†
Radiological variables			
Modic I and II			
Not present	21/131	52/62	0.76†
Modic primary type I present	48/104	65/56	
Modic primary type II present	55/97	56/60	
Modic I and II present	28/124	57/59	
Modic CC[e]	55/97	69/53	0.05†
Disc height reduction[f]	104/48	61/54	0.46†
Nucleus pulposus grade 3/4	133/19	58/63	0.66†
Facet joint arthropathy grade 2/3	14/138	64/58	0.65†
Posterior HIZ	63/89	54/62	0.33†

Continuous and categorical variables

SF-36 scores range from 0 to 100, with higher scores corresponding to better health status

Waddell's FABQ scale ranges from 0 to 24 (physical) and from 0 to 42 (work), with lower scores indicating less severe symptoms

HSCL-25 = Hopkins Symptom checklist; HSCL-25 reflects emotional distress, and scores range from 1 to 4, with lower scores indicating less severe symptoms

P value; Indicating if the baseline variable is associated with working or not at follow-up

FABQ Fear of Avoidance Belief Questionnaire

* Independent two-sided *t* test; †Pearson's χ^2; ‡Mann–Whitney *U* test; ¶Fischer's exact test

[a] Classified according to socioeconomic status from Statistics Norway consisting of six groups: manual low, manual high, routine nonmanual low, routine nonmanual high, professional low or professional high. Because there were few patients in each group, unskilled and skilled workers were merged into the same group and nonmanual in one group. Consequently, two groups were analyzed, manual and nonmanual[14]

[b] Because there were fewer than five subjects in the category of education at university level, they were merged with 13-15 years of education (college)

[c] Working versus not working;Including part-time work as working

[d] Uden et al. [19]

[e] More than 50% of the craniocaudal diameter

[f] More or less than 40% height reduction (more = yes)

The Oswestry Disability Index (ODI) ranges from 0 to 100, with lower scores indicating less severe symptoms

Back pain was calculated using a horizontal scale ranging from 0 (no pain) to 100 (worst pain imaginable) with word anchors at the beginning and end

predictors found in previous studies [7, 27, 28]. Zindrick et al. [10] concluded in a review that there was no definitive evidence about which variables that affect outcome in disc prosthesis surgery. Siepe et al. [28, 29], reported better outcomes for disc surgery in younger patients, for monosegmental surgery (especially at the L4/L5 level) and in patients with lower disc height. Bertagnoli et al. found better outcomes in patients with disc height >4 mm and no facet joint arthritis, and Guyer et al. found that time off work before treatment predicted outcome [30, 31]. We could not confirm these findings.

Table 4 Step 3 and final model in the multiple logistic regression model

	Final model (step 3)			
	P value	*B*	OR	95% CI for OR
Surgery ($n = 88$)				
Physical variables				
Duration of back pain (5 years)[a]	0.01	0.7	1.9	1.2–3.2
Psychological variables				
FABQ-work (5 points)[b]	0.007	0.5	1.7	1.2–2.4
Radiological variables				
Modic I or II[c]	0.04	1.7	5.3	1.1–25.3
Rehabilitation ($n = 66$)				
Physical variables				
Daily consumption of narcotics	0.02	2.5	23.6	2.1–266.8
ODI (5 points)[d]	0.002	1.0	2.5	1.4–4.5
Psychological variables				
HSCL (0.5 points)[e]	0.02	0.9	2.4	1.1–5.1

Baseline variables associated with 15 ODI points improvement at follow-up in the surgical and nonsurgical groups. Continuous and categorical variables

B regression coefficient, *OR* Odds ratio

[a] OR for change of 5 years

[b] Waddell's FABQ work scale ranges from 0 to 42, with lower scores indicating less severe symptoms; OR for change of 5 points

[c] Modic changes compared with not present

[d] The Oswestry Disability Index (ODI) ranges from 0 to 100, with lower scores indicating less severe symptoms; OR for change of 5 points

[e] HSCL-25 = Hopkins Symptom checklist; HSCL-25 reflects emotional distress, and scores range from 1 to 4, with lower scores indicating less severe symptoms; OR for change of 0.5 points

Table 5 Step 3 and final model in the multiple logistic regression model

Surgery and rehabilitation ($n = 154$)	Final model (step 3)			
	P value	*B*	OR	95% CI for OR
Sociodemographic variables				
Working[a]	0.02	1.4	4.1	1.2–13.2
Psychological variables				
FABQ work[b]	0.03	0.2	1.3	1.0–1.5

Baseline variables associated with working at follow-up in the surgical and nonsurgical groups combined. Continuous and categorical variables

OR Odds ratio, *B* regression coefficient

[a] Working versus not working. Including part-time work as working

[b] Waddell's FABQ work scale ranges from 0 to 42, with lower scores indicating less severe symptoms; OR for change of 5 points

In the present study, we found that for patients treated with disc prosthesis, long duration of back pain and high FABQ-W score at baseline were significantly associated with worse outcome assessed by ODI, and Modic primary or secondary type I and/or II were significantly correlated with better outcome. Furthermore, supported by a sub analysis, Modic primary type I increased the chance of being among the 15 patients with best results after surgery ($P = 0.008$; OR = 10.1, CI 1.8–56.0). No former studies have identified Modic changes as predictor for outcome after treatment with disc prosthesis, but one former study reported that patients with chronic LBP, degenerative discs and Modic type I undergoing fusion, had favorable outcome [27]. Our findings suggest that removing the disc, possibly causing the pain in such patients, may predict good outcome.

Patients with long-lasting back pain and high FABQ-W had a less favorable outcome. These two variables represent psychosocial aspects of chronic back pain, thus social interactions with adaptation to the sick role or a negative attitude and belief toward possible recovery may influence the prognosis. Central sensitization can also contribute to the maintenance of pain and disability in patients with long-lasting low back pain [32].

We found different predictors for success after rehabilitation. High baseline ODI, low HSCL-25 and not using narcotics predicted better results. High disability and pain scores have also been found to predict a larger improvement in functional outcome after rehabilitation in former studies [7]. This could not only represent regression to the mean or ceiling effects, but could also influence the treatment effect per se. In our randomized study, there was a difference between groups in favor of surgery, but as illustrated in Fig. 1, patients in the surgical group and rehabilitation group with high baseline ODI (>55) experienced similar results, indicating that especially these patients should be treated with rehabilitation before considering surgery [12]. In a review, Van der Hulst et al. [7] could not find a generic set of predictors of outcome for patients with LBP treated with multidisciplinary rehabilitation. However, several work-related parameters, coping style and pain intensity were associated with outcome. In the present study, patients reporting high emotional distress had poorer outcome which suggest that parts of the cognitive treatment did not succeed in these patients. Brage et al. [33] reported that patients with LBP and high emotional distress had an increased risk of disability, emphasizing the importance of treating this condition. In prior studies, multidisciplinary rehabilitation incorporating analgesic medication withdrawal has been associated with clinical improvement of pain intensity and functioning [34]. We found worse outcomes in patients using narcotics daily, which might indicate that the rehabilitation program

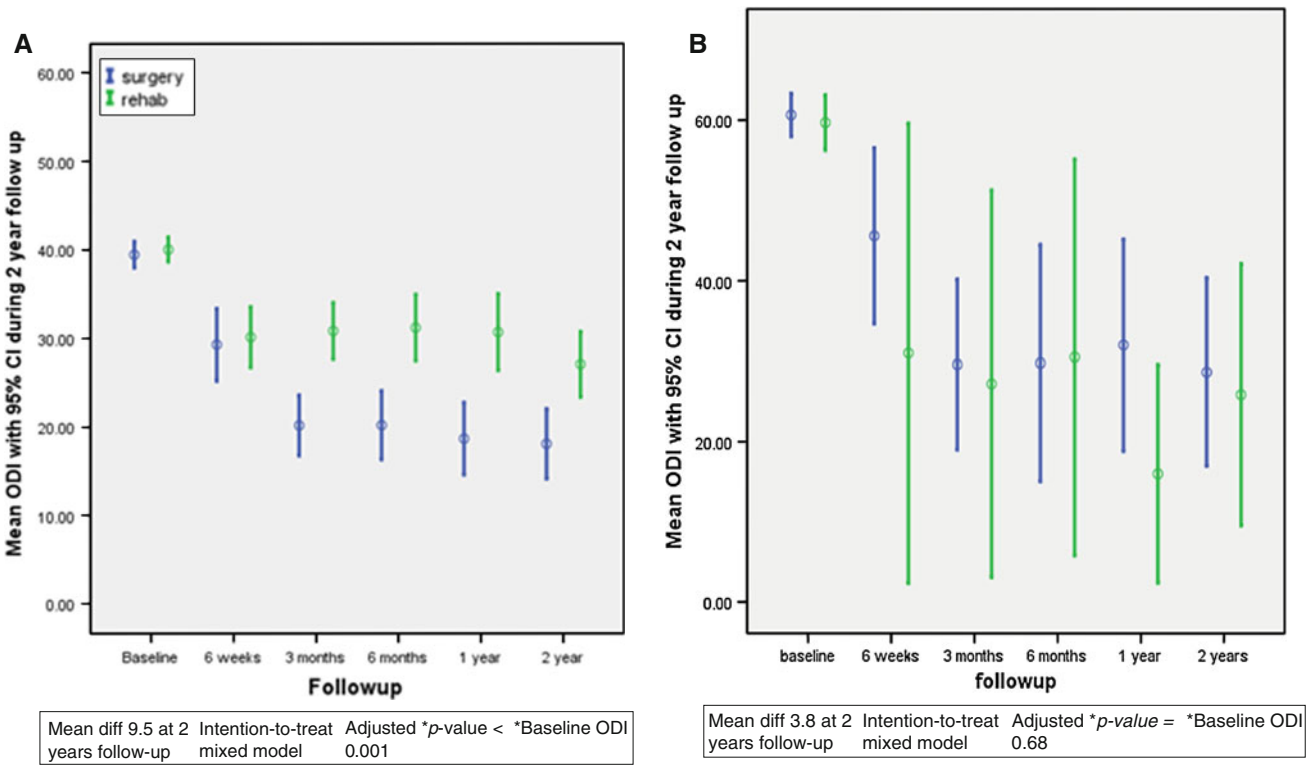

Fig. 1 The Oswestry Disability Index (ODI) ranges from 0 to 100, with lower scores indicating less severe symptoms. **a** ODI during follow-up of patients with ODI at baseline <55 points, **b** ODI during follow-up of patients with ODI at baseline ≥55 points

failed both in patients with high HSCL-25 and daily usage of narcotics. However, the confidence interval for daily usage of narcotics was very large and OR may be overestimated in small sample sizes [35]. In a recent systematic review, opioid therapy was not associated with reduction of pain compared with placebo [36]. This indicates that opioid treatment of patients with chronic LBP should be avoided.

We found being at work and low FABQ-W at baseline to predict being at work at follow-up in both groups. This is in line with previous studies [7]. High FABQ-W is identified as a predictor for future sick listing or disability in patients with acute and sub-acute LBP [37].

The main strength of our study is the stringent inclusion and exclusion criteria making the patient population quite homogenous. Furthermore, both surgical and rehabilitation treatments were standardized across study centres. We did a multivariate analysis to evaluate confounding effects. A multivariate logistic analysis probably reduces the chance of detecting nonclinical significant predictors and may also make the result more clinically useful compared with a multivariate linear regression analysis.

Some limitations should be considered. First, having several baseline variables increases the risk of making a type I error. A larger study cohort would increase the power of the study. Furthermore, the choice of outcome variables can influence results. An improvement of 15 ODI points has been evaluated as clinically important. Electing a 50% reduction in ODI at follow up as an outcome variable might give a larger reduction on average than 15 ODI points and perhaps with different predictors. Furthermore, our patient population was highly selective and the results are strictly applicable in patients with localized LBP and moderate to severe degenerative discs at the two lower lumbar levels. The two investigated cohorts are similar, but since the analysis is done "as treated", a strict comparison of predictors between treatments should be avoided (Table 1).

As former studies and reviews have stated, the patient populations in studies within this field are heterogeneous, methods of analysis are different and outcome measures are various. We could confirm some of the variables as predictors, but there is no consistent set of variables in the literature upon which to rely, when choosing patients for total disc replacement or rehabilitation. However, many of the variables seemed to be associated to some degree. This complexity may also reflect the fact that the variables together are of major importance, but individually have only small influence. It is probably necessary to merge studies in a meta-analysis to gather more information about which variables are predictors.

It has not been shown previously that FABQ-W predicts outcome in LBP patients undergoing disc surgery. We suggest that questionnaires including FABQ-W, ODI and measures of emotional distress such as HSCL-25 should be included in the pre-treatment evaluation of patients treated for degenerative disc disease. It could also be argued that patients should be treated earlier than is current practice to reduce the development of chronicity with therapeutic resistance. It is also reasonable that all patients and especially patients with high ODI and high FABQ-W should be treated with multidisciplinary rehabilitation before disc replacement becomes an option. Patients with low FABQ-W or Modic changes type I or II should be considered for surgery if rehabilitation fails. However, our findings need to be confirmed in future studies.

Acknowledgments We want to thank the patients participated in the study, the South Eastern Norway Regional Health Authority and EXTRA funds from the Norwegian Foundation for Health and Rehabilitation, through the Norwegian Back Pain Association, for financial support, The Coast Hospital for Physical Medicine and Rehabilitation, Stavern, for videos and material for instructions on the rehabilitation intervention, Hege Andresen at St.Olavs Hospital, Trondheim, for data coordination, Astrid Woodhouse and Kirsti Vanvik from St.Olavs Hospital for performing the 2-year control and Prof. Leiv Sandvik for statistical counseling and support. Financial support was received from the South Eastern Norway Regional Health Authority and EXTRA funds from the Norwegian Foundation for Health and Rehabilitation, through the Norwegian Back Pain Association.

Conflict of interest All authors involved declare that they have no conflict of interests and no financial disclosures to report.

Appendix

From University Hospital North Norway, Tromsø (included $n = 8$ patients): Department of Orthopedic Surgery: Odd-Inge Solem, MD, Department of Neurosurgery: Jens Munch-Ellingsen, MD, PhD, and from Department of Physical Medicine and Rehabilitation Franz Hintringer, MD, Anita Dimmen Johansen, ergonomist, Guro Kjos, PT.

From Trondheim University Hospital, Trondheim (included $n = 21$ patients): National centre for spinal disorders, Department of Neurosurgery: Øystein P Nygaard, PhD, Hege Andresen, RN, Helge Rønningen, MD, Professor, Kjell Arne Kvistad, MD, PhD, and from Multidiscipline spinal unit, Dep. of Physical Medicine and Rehabilitation; Bjørn Skogstad, MD, Janne Birgitte Børke, PT, MSc, Erik Nordtvedt, PT, Magne Rø, MD, Gunnar Leivseth, MD, Professor.

From Haukeland University Hospital, Bergen (included $n = 64$ patients): Kysthospitalet in Hagevik, Department of Orthopedic Surgery: Sjur Braaten, MD, Turid Rognsvåg, PT, MSc, Gunn Odil Hirth Moberg, secretary, and From The Outpatient Spine Clinic. Department of Physical Medicine and Rehabilitation: Lars Geir Larsen, PT, Vibeche Iversen, RN, Ellen H Haldorsen, cand, psychol, PhD, Elin Karin Johnsen, RN, Kristin Hannestad, PT;

From Stavanger University Hospital, Stavanger (included $n = 27$ patients): Department of Orthopedic Surgery: Endre Refsdal, MD.

From Oslo University Hospital, Oslo (included $n = 53$ patients): Department of Orthopaedics: Vegard Slettemoen, RN, Kenneth Nilsen, RN, Kjersti Sunde, RN, Helenè E Skaara, PT, MSc, and from Department of Physical Medicine and Rehabilitation: Anne Keller, MD, PhD, Berit Johannessen, PT, Anna Maria Eriksdotter, PT, MSc.

References

1. van den Eerenbeemt KD, Ostelo RW, van Royen BJ, Peul WC, van Tulder MW (2010) Total disc replacement surgery for symptomatic degenerative lumbar disc disease: a systematic review of the literature. Eur Spine J 19:1262–1280
2. Blumenthal S, McAfee PC, Guyer RD, Hochschuler SH, Geisler FH, Holt RT, Garcia R Jr, Regan JJ, Ohnmeiss DD (2005) A prospective, randomized, multicenter Food and Drug Administration investigational device exemptions study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part I: evaluation of clinical outcomes.[erratum appears in Spine. 2005 Oct 15;30(20):2356]. Spine 30:1565–1575
3. Zigler J, Delamarter R, Spivak JM, Linovitz RJ, Danielson GO III, Haider TT, Cammisa F, Zuchermann J, Balderston R, Kitchel S, Foley K, Watkins R, Bradford D, Yue J, Yuan H, Herkowitz H, Geiger D, Bendo J, Peppers T, Sachs B, Girardi F, Kropf M, Goldstein J (2007) Results of the prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement versus circumferential fusion for the treatment of 1-level degenerative disc disease. Spine 32:1155–1162
4. Berg S, Tullberg T, Branth B, Olerud C, Tropp H (2009) Total disc replacement compared to lumbar fusion: a randomised controlled trial with 2-year follow-up. Eur Spine J 18:1512–1519
5. Fritzell P, Berg S, Borgstrom F, Tullberg T, Tropp H (2011) Cost effectiveness of disc prosthesis versus lumbar fusion in patients with chronic low back pain: randomized controlled trial with 2-year follow-up. Eur Spine J 20:1001–1011
6. Brox JI, Sorensen R, Friis A, Nygaard O, Indahl A, Keller A, Ingebrigtsen T, Eriksen HR, Holm I, Koller AK, Riise R, Reikeras O (2003) Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. Spine 28:1913–1921
7. van der Hulst M, Vollenbroek-Hutten MM, Ijzerman MJ (2005) A systematic review of sociodemographic, physical, and psychological predictors of multidisciplinary rehabilitation-or, back school treatment outcome in patients with chronic low back pain. Spine 30:813–825
8. Macedo LG, Maher CG, Latimer J, McAuley JH (2009) Motor control exercise for persistent, nonspecific low back pain: a systematic review. Phys Ther 89:9–25
9. van Middelkoop M, Rubinstein SM, Kuijpers T, Verhagen AP, Ostelo R, Koes BW, van Tulder MW (2011) A systematic review

on the effectiveness of physical and rehabilitation interventions for chronic non-specific low back pain. Eur Spine J 20:19–39
10. Zindrick MR, Tzermiadianos MN, Voronov LI, Lorenz M, Hadjipavlou A (2008) An evidence-based medicine approach in determining factors that may affect outcome in lumbar total disc replacement. Spine 33:1262–1269
11. Hagg O, Fritzell P, Ekselius L, Nordwall A (2003) Predictors of outcome in fusion surgery for chronic low back pain. A report from the Swedish Lumbar Spine Study. Eur Spine J 12:22–33
12. Hellum C, Johnsen LG, Storheim K, Nygaard OP, Brox JI, Rossvoll I, Ro M, Sandvik L, Grundnes O (2011) Surgery with disc prosthesis versus rehabilitation in patients with low back pain and degenerative disc: two year follow-up of randomised study. BMJ 342:d2786
13. Fairbank JCTM, Pynsent PBP (2000) The Oswestry Disability Index. Spine 25:2940–2953 (review)
14. Statistics Norway (1984) Standard Classifications of Socioeconomic Status
15. Statistics Norway (1998) Standard Classification of Occupations
16. Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L (1974) The Hopkins symptom checklist (HSCL): a self-report symptom inventory. Behav Sci 19:1–15
17. Waddell G, Newton M, Henderson I, Somerville D, Main CJ (1993) A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain 52:157–168
18. Ware JE Jr, Sherbourne CD (1992) The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 30:473–483
19. Uden A, Astrom M, Bergenudd H (1988) Pain drawings in chronic back pain. Spine 13:389–392
20. Modic MT, Steinberg PM, Ross JS, Masaryk TJ, Carter JR (1988) Degenerative disk disease: assessment of changes in vertebral body marrow with MR imaging. Radiology 166:193–199
21. Fujiwara A, Tamai K, Yamato M, An HS, Yoshida H, Saotome K, Kurihashi A (1999) The relationship between facet joint osteoarthritis and disc degeneration of the lumbar spine: an MRI study. Eur Spine J 8:396–401
22. Aprill C, Bogduk N (1992) High-intensity zone: a diagnostic sign of painful lumbar disc on magnetic resonance imaging. Br J Radiol 65:361–369
23. Luoma K, Riihimaki H, Luukkonen R, Raininko R, Viikari-Juntura E, Lamminen A (2000) Low back pain in relation to lumbar disc degeneration. Spine 25:487–492
24. Jensen TS, Sorensen JS, Kjaer P (2007) Intra- and interobserver reproducibility of vertebral endplate signal (modic) changes in the lumbar spine: the Nordic Modic Consensus Group classification. Acta Radiol 48:748–754
25. Masharawi Y, Kjaer P, Bendix T, Manniche C, Wedderkopp N, Sorensen JS, Peled N, Jensen TS (2008) The reproducibility of quantitative measurements in lumbar magnetic resonance imaging of children from the general population. Spine 33:2094–2100
26. Hosmer DW, Lemeshow S, May S (2008) Applied Survival Analysis. Regression Modelilng of Time-to-Event Data. John Wiley & Sons, Inc, Hoboken
27. Esposito P, Pinheiro-Franco JL, Froelich S, Maitrot D (2006) Predictive value of MRI vertebral end-plate signal changes (Modic) on outcome of surgically treated degenerative disc disease. Results of a cohort study including 60 patients. Neurochirurgie 52:315–322
28. Siepe CJ, Mayer HM, Heinz-Leisenheimer M, Korge A (2007) Total lumbar disc replacement: different results for different levels. Spine 32:782–790
29. Siepe CJ, Hitzl W, Meschede P, Sharma AK, Khattab MF, Mayer MH (2009) Interdependence between disc space height, range of motion and clinical outcome in total lumbar disc replacement. Spine (Phila Pa 1976) 34:904–916
30. Guyer RD, Siddiqui S, Zigler JE, Ohnmeiss DD, Blumenthal SL, Sachs BL, Hochschuler SH, Rashbaum RF (2008) Lumbar spinal arthroplasty: analysis of one center's twenty best and twenty worst clinical outcomes. Spine 33:2566–2569
31. Bertagnoli R, Kumar S (2002) Indications for full prosthetic disc arthroplasty: a correlation of clinical outcome against a variety of indications. Eur Spine J 11(Suppl 2):S131–S136
32. Campbell JN, Meyer RA (2006) Mechanisms of neuropathic pain. Neuron 52:77–92
33. Brage S, Sandanger I, Nygard JF (2007) Emotional distress as a predictor for low back disability: a prospective 12-year population-based study. Spine (Phila Pa 1976) 32:269–274
34. Crisostomo RA, Schmidt JE, Hooten WM, Kerkvliet JL, Townsend CO, Bruce BK (2008) Withdrawal of analgesic medication for chronic low-back pain patients: improvement in outcomes of multidisciplinary rehabilitation regardless of surgical history. Am J Phys Med Rehabil 87:527–536
35. Nemes S, Jonasson JM, Genell A, Steineck G (2009) Bias in odds ratios by logistic regression modelling and sample size. BMC Med Res Methodol 9:56
36. Martell BA, O'Connor PG, Kerns RD, Becker WC, Morales KH, Kosten TR, Fiellin DA (2007) Systematic review: opioid treatment for chronic back pain: prevalence, efficacy, and association with addiction. Ann Intern Med 146:116–127
37. Linton SJ, Hallden K (1998) Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain. Clin J Pain 14:209–215