Lumbar total disc replacement: Predictors for long-term outcome

Håvard Furunes, Christian Hellum, Jens Ivar Brox, Ivar Rossvoll, Ansgar Espeland, Linda Berg, Helga Maria Brøgger, Milada Cvancarova Småstuen, Kjersti Storheim.

Abstract

Purpose

We aimed to identify patient characteristics associated with favourable long-term outcomes after lumbar total disc replacement (TDR).

Methods

We analysed a cohort of 82 patients with degenerative disc and chronic low back pain (LBP) who were treated with TDR and originally participated in a randomised trial comparing TDR and multidisciplinary rehabilitation. Potential predictors were measured at baseline, and the outcomes assessed eight years after they received allocated treatment. Outcome measures were dichotomised according to whether the participants achieved a clinically important functional improvement (15 points or more on the Oswestry Disability Index, ODI) (primary outcome) and whether they were employed at eight-year follow-up (secondary outcome). Associations between potential predictors and outcomes were modelled using logistic regression. For the secondary outcome, the results were also organised in a prediction matrix and expressed as probabilities.

Results

For 71 patients treated with TDR according to protocol, the follow-up time was eight years. For a subgroup of 11 patients randomised to rehabilitation who crossed over and received TDR, the median postoperative follow-up time was 72 (range 41-88) months. Of all assessed baseline variables, only presence of Modic changes (type 1 and/or 2) was statistically significantly associated with an improvement of ≥ 15 ODI points. The probability of employment at eight-year follow-up was 1 % for patients with ≥ 1 year of sick leave, comorbidity, ODI ≥ 50 and ≤ nine years of education prior to treatment, and 87 % for patients with < 1 year of sick leave, no comorbidity, ODI < 50 and higher education.

Conclusions

Patients with Modic changes prior to the TDR surgery were more likely to report a clinically important functional improvement at long-term follow-up. Comorbidity, low level of education, long-term sick leave and high ODI score at baseline were associated with unemployment at long-term follow-up.

Keywords: Low back pain, degenerative disc, lumbar total disc replacement, patient selection

Introduction

Total disc replacement (TDR) is a surgical option for selected patients with low back pain (LBP) and degenerative intervertebral disc when non-operative treatment fails. Despite promising short-term results, the authors of a Cochrane report [1] encourage spine surgeons to be cautious about implementing the surgical procedure on a large scale because complications may arise after several years. This view is supported by a recent systematic review [2] comparing TDR and spine fusion. Over the last years, a few studies with long-term follow-up

after TDR surgery have been published [3-8]. A clinically important improvement according to FDA criteria [5] (15 points improvement or more on the Oswestry Disability Index (ODI)) is reported by 68-87 % of patients 5-8 years after TDR [5, 6, 8], and 67-88 % of patients are employed at follow-up 5-13 years after TDR [3, 4, 6, 7].

Park et al. [9] showed inferior long-term results of TDR in patients that were presumed to be bad candidates for the procedure compared to patients that were presumed to be good candidates. The categorisation was based on the presence or absence of suggested contraindications for TDR (surgery at the adjacent level of a fused segment, spondylolisthesis, facet joint arthritis and lateral recess stenosis). In the randomised trial from which our data are extracted, 24 % of the patients had no symptoms of back pain eight years after TDR, and yet 8 % described themselves as "worse than ever" [10]. This illustrates the obvious need for improved patient selection criteria for disc replacement.

At two-year follow-up, Hellum et al. [11] found that the best predictors for a clinically important improvement (≥ 15 ODI points) after TDR were short preoperative duration of LBP, low Fear-Avoidance Beliefs about work (FABQ-work) and the presence of Modic changes at baseline. In the only study examining the association between baseline characteristics and mid- to long-term outcome, Gornet et al. [12] found that better clinical outcome at five-year follow-up was related to higher grades of degeneration of the index level before surgery. Still, these reports provide limited information about patient characteristics associated with the long-term outcome after TDR.

The aim of this study was to identify baseline characteristics associated with a clinically important improvement (≥ 15 ODI points) (primary outcome) and with employment (secondary outcome) at eight-year follow-up after inclusion in this prospective study.

Methods

Study design

This is a prospective cohort study of patients treated with TDR for chronic LBP and degenerative intervertebral lumbar disc. The patients were included in a multicentre randomised trial comparing TDR with multidisciplinary rehabilitation [13], and data are extracted from the eight-year follow-up.

Ethical concerns

The eight-year follow-up of the randomised trial was approved by the Norwegian Regional Ethical Committee–South-East C (2011/2177). The project was registered at www.clinicaltrial.gov under the identifier NCT01704677 before it commenced in accordance with the Helsinki Declaration and the ICH-GCP guidelines.

Results are reported according to the STROBE standard for reporting cohort studies.

Participants

Inclusion criteria for the original randomised trial were age 25-55 years, LBP as the main symptom for at least one year, ODI score ≥ 30, conservative treatment for ≥ six months without sufficient effect and degenerative changes in the intervertebral disc L4/L5 and/or L5/S1. For further details see Hellum et al. [13]. The patients included in the present cohort study were either treated with TDR according to the randomisation, or they crossed over from the rehabilitation group and were treated with TDR. We did not exclude patients who had been reoperated or had received additional non-operative treatment.

Study intervention

The patients were treated with a surgical procedure in which the degenerative intervertebral lumbar disc was removed and replaced with an artificial disc (ProDisc II, Synthes Spine). The

treatment took place at one of the five Norwegian University Hospitals where the study was conducted. A more detailed description of the TDR procedure has been reported previously [13].

Outcome measures (dependent variables)

The primary outcome measure was change in self-reported physical function from baseline to eight-year follow-up, measured by the ODI [14]. Change in ODI was dichotomised, and an improvement of ≥ 15 points was categorised as a minimal clinically important improvement, according to FDA criteria [5]. The secondary outcome measure was self-reported work status at eight-year follow-up. Patients who reported full- or part-time employment, or were students, were categorised as employed.

Potential predictors of outcome (independent variables)

Variables tested for predictive value were collected at baseline and categorised as socio-demographic, clinical, psychological variables and pain, and radiological variables (Table 1).

Socio-demographic variables

All socio-demographic variables were patient reported. Patients were categorised as manual or non-manual workers according to the Norwegian Standard Classification of Socioeconomic Status [15]. The classification consists of six groups, but since there were few patients in each group, they were dichotomised as manual or non-manual workers. Educational level was categorised according to the International Standard of Classification of Education (IECED) [16]. Work status was categorised as employed (part time or full time) or unemployed. In addition, information on duration of sick leave, smoking, gender and age was collected.

Clinical variables

Clinical variables included prior discectomy, level(s) operated on with TDR, presence of comorbidity, ODI and body mass index (BMI). The predicting value of a threshold level in

baseline ODI of 55 points has been tested previously [11]. Since there were too few patients with an ODI ≥ 55 points at baseline in the present sample, we chose to test a threshold level of 50 points. The variables were patient reported, except level(s) operated on, which was reported by the surgeon.

Psychological variables and pain

Psychological variables were Hopkins Symptom Check List (HSCL-25) [17], Fear-Avoidance Belief Questionnaire (FABQ) [18] and the Mental Component Scale (MCS) part of SF-36 [19]. Pain variables were LBP intensity (Visual Analogue Scale, VAS), pain drawing categorised as pain below the waist or pain above the waist (with or without pain below the waist) [20], duration of LBP and daily consumption of narcotics (yes / no).

Radiological variables

Pelvic incidence [21] was measured on radiographs obtained at the last follow-up by an experienced radiologist blinded to the clinical data, and was analysed as a baseline variable since it describes the fixed relationship between the femoral heads and the endplate of the sacrum – which should remain unaltered after TDR. Pelvic incidence was dichotomised as < / ≥ 55, as recommended by Prof. Le Huec (personal communication). All other radiological variables (Modic changes [22], disc height reduction [23], nucleus pulposus grade [24], facet arthropathy [25] and posterior high intensity zone [26]) were evaluated independently on pre-treatment images by three experienced radiologists blinded to the 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) [27].

Statistical analysis

Continuous variables were described as medians and ranges, categorical variables as proportions and percentages. Outcome variables (clinical improvement (yes / no) and

employment (yes / no)) were modelled as the dependent variables and selected baseline covariates as the independent variables. Possible associations between selected variables and outcomes were modelled using binary logistic regression. Potential predictors that were highly associated with each other were excluded to avoid multicollinearity. Due to a limited sample size and few patients who improved / were employed, we fit models with a maximum of four covariates to avoid overfitting. Therefore, only baseline characteristics that were statistically significantly ($p < 0.05$) associated with the outcome in univariate analyses were entered into the final multiple model. Further, the results from the multiple model were used to compute probabilities for the outcome given any selected value of the covariates, and the probabilities were expressed in a prediction matrix. The results were expressed as odds ratios (OR) with 95 % confidence intervals (CI). Since the sample size was limited, we were not able to set aside a test set for validation, and instead performed a leave-one-out cross-validation [28]. A sensitivity analysis was performed, excluding patients who were originally randomised to rehabilitation and patients who had received additional spinal surgery after the TDR. All tests were two-sided and p-values < 0.05 were considered statistically significant. Since our study was exploratory, no correction for multiple testing was performed. The statistical analyses were performed with SPSS version 24.0.

Results

Of the 86 patients randomised to surgery, nine did not receive the surgical treatment and nine were lost to follow-up (five lost contact, four withdrew consent). Hence, 71 patients were analysed eight years postoperatively. In addition, we included 14 patients randomised to rehabilitation who crossed over and were treated with TDR. Of these, 11 were available for follow-up (median time since surgery was 72 (range 41-88) months). Consequently, 82 patients (82 %) were included in the final cohort analyses (Figure 1). Nine of these 82 patients (11 %) had been reoperated (one because of implant dislocation, one with neurostimulator

implantation, two with spinal fusion and five with decompression of spinal stenosis). Median time since reoperation was 37 (range 1-103) months.

Overall, 52 patients (63 %) achieved a clinically important improvement of ≥ 15 ODI points, and 42 patients (51 %) were employed eight years after they were included in the study.

Baseline variables significantly associated with the clinically important improvement were the presence of Modic changes (type 1 and/or 2) (OR 5.0, 95 % CI 1.4-18.2, p=0.01) and the extent of Modic changes (> 50 % of vertebral body height) (OR 3.8, 95 % CI 1.3-11.5, p=0.02) (Table 2). However, the presence of Modic changes and the extent of Modic changes were significantly associated with each other (p=0.01) and could not be included in the same model. Therefore, we did not proceed with the fitting of a prediction model.

Baseline variables significantly associated with the status of being employed at eight-year follow-up were < 12 months of sick leave before treatment (OR 4.1, 95 % CI 1.6-10.6, p=0.003), absence of comorbidity (OR 4.4, 95 % CI 1.4-13.8, p=0.01), ODI < 50 points (OR 3.6, 95 % CI 1.0-12.5) and high level of education (> nine years) (OR 3.6, 95 % CI 1.1-11.2, p=0.03) (Table 3). In addition, FABQ-work was statistically significantly associated with employment at eight-year follow-up (OR 0.9, 95 % CI 0.9-1.0, p=0.01). However, in the multivariate analysis with comorbidity, education level, ODI ≥ 50 and ≥ 12 months' sick leave, including FABQ-work weakened the predictive power of the model, and we therefore did not include FABQ-work in the final multiple model (Table 3). We found significant differences in the probabilities of being employed corresponding to the different combinations of the baseline variables. The probability of employment at the last follow-up was 1 % (95 % CI 0-4 %) for patients with ≥ 12 months' sick leave, comorbidity, ODI ≥ 50 and ≤ nine years of education prior to treatment, and 87 % (95 % CI 80-94 %) for patients with < 12 months' sick leave, no comorbidity, ODI < 50 and higher education (Figure 2).

Sensitivity analyses confirmed our results. When we excluded patients who were reoperated or who had crossed over from the rehabilitation group, the presence of Modic changes at baseline was still the only baseline variable that was significantly associated with a clinically important improvement (≥ 15 ODI points) (OR 6.5, 95 % CI 1.4-30.0, p=0.02). Baseline characteristics significantly associated with employment after eight years were still <12 months of sick leave before treatment (OR 3.6, 95 % CI 1.3-10.0, p=0.01), absence of comorbidity (OR 4.7, 95 % CI 1.3-16.6, p=0.02), ODI < 50 (OR 4.9, 95 % CI 1.2-19.9, p=0.02), higher education (OR 4.1, 95 % CI 1.2-14.6, p=0.01) and FABQ-work (OR 1.1, 95 % CI 1.0-1.1, p=0.01).

Discussion

In this prospective cohort study, the presence of Modic changes (type 1 and/or 2) was statistically significantly associated with a clinically important improvement (≥ 15 ODI points). Patients with a shorter duration of sick leave, absence of comorbidity, lower ODI score and higher education were more likely to be employed at eight-year follow-up.

The extent of Modic changes (> 50 % of the vertebral body height) was significantly associated with both the presence of Modic changes and the outcome (≥ 15 points improvement in ODI score). Therefore, the extent of Modic changes may be as important as the presence of Modic changes in regards to the association with the outcome.

The positive association between Modic changes and ≥ 15 points improvement in ODI score after TDR in our study should be interpreted in light of the findings in a recent systematic review on the impact of Modic changes on outcome after lumbar spine surgery [29]. This review identified four TDR studies (including the two-year results from the present study [13]). One study found no association between Modic changes and ODI or LBP after TDR, and the remaining three had conflicting findings about which types of Modic changes (type 1,

type 2, or both types combined) were related to ODI or pain after TDR. Although Modic changes seem to be associated with improved outcome after TDR, the association is not consistent between different studies or outcomes, and it should be examined in larger high-quality studies.

Gornet et al. [12] found significantly less improvement in ODI score at two- and five-year follow-up after TDR in patients with workers' compensation. They also found a statistically significant association between a favourable outcome measured with ODI at five-year follow-up and higher grades of disc degeneration preoperatively, presence of Modic type 2 changes and a smaller proportion of the overall lumbar lordosis (L1-S1) at the treatment level.

Shorter duration of sick leave, absence of comorbidity, lower ODI score and higher education at baseline increased the probability of employment at eight-year follow-up in our prediction matrix. These findings are plausible, but in the literature there is no consensus on baseline characteristics that predict return to work after surgery in patients with chronic LBP. In populations including mostly non-operated patients with LBP or sciatica, Cougot et al. [30] found that the patient's profession was the only predictor for return to work in health care workers with LBP. In patients with sciatica, Grøvle et al. [31] found that lower age, better general health, lower baseline sciatica bothersomeness, lower score on the FABQ-work and a negative straight leg raising test result were significantly associated with a higher probability of returning to work. McGirth et al. [32] found that preoperative depression, arthritis and prolonged preoperative opioid use reduced the likelihood of returning to work in patients labeled as having degenerative chronic LBP without workers' compensation. In a longitudinal study of women, Nordeman et al. [33] found that the six-minute walk test, depression and earlier ability to work predicted the ability to work at two-year follow-up. Hence, the biopsychosocial factors at baseline associated with employment at follow-up in our study find broad support in the literature.

The strengths of this study are the prospective design, substantial follow-up rate (82 %), long follow-up time, biopsychosocial approach and public financing.

The study also had limitations. First, a minimal clinically important change (MCIC) could be defined in several ways. We define a clinically important improvement as 15 points improvement in ODI score from baseline, in agreement with FDA studies [5, 8] and a previous report from the present study [11]. A clinically important improvement is also commonly defined as a 30 % improvement on ODI [1], and in the two-year follow-up in the randomised study from which our data are extracted, the clinically important improvement was calculated as 12.88 ODI points based on Receiver Operator Curve (ROC) analysis [34]. An ODI score ≤ 22 after surgery for degenerative disorders of the lumbar spine is suggested as a threshold for a "satisfactory symptom state", regardless of the baseline score [35]. Different outcome measures may be associated with different baseline variables.

Secondly, the sample size is limited. A larger simple size would have allowed us to fit a larger prediction model, perform a validation and possibly identify further variables associated with the outcome.

The cut-off values of the independent variables represent a third limitation. In order to create a prediction matrix that could help clinicians and patients choose the right treatment for chronic LBP, the independent variables had to be dichotomised. Due to the limited sample size, the cut-off values were not only based on clinical recommendations, but also on statistical properties that gave the best separation among subgroups of patients. The associations might have been weakened if we had used other cut-off values for the independent variables.

A fourth limitation is the relatively strict selection of patients. Our findings may not apply to the general population with chronic LBP. On the other hand, TDR is only indicated in

selected patients, and we believe that the participants of this study are representative as candidates for TDR.

Fifthly, we have limited knowledge of the natural course of chronic LBP over eight years. However, Peng et al. [36] observed a small and clinically unimportant improvement from 46.4 to 44.0 points on ODI over four years in an observational study of patients with chronic LBP. Therefore, we may assume that the change in physical function in our cohort is mainly caused by the intervention, and only minimally influenced by the natural course of LBP.

Further, the substantial number of patients who had received treatments other than TDR might have influenced the long-term results. Nine patients were reoperated. Patients who undergo reoperations generally have inferior results [10, 37], which may weaken the association between baseline characteristics and a clinically important improvement. Moreover, the 11 patients who crossed over from the rehabilitation group to TDR had a shorter observation time. However, the sensitivity analysis that excluded those who were reoperated and those who crossed over from rehabilitation showed results similar to those of the main analysis.

In conclusion, the presence of Modic changes was statistically significantly associated with long-term improvement after TDR. Moreover, our visual prediction matrix, combining readily available patient characteristics, revealed substantial differences between patient groups regarding the probability of employment at long-term follow-up. The prediction matrix might help to improve the patient selection for TDR, and act as a guide for physicians and patients choosing a treatment for chronic LBP.

Funding

The study was funded by Oslo University Hospital, South Eastern Norway Regional Health Authority, and EXTRA funds from the Norwegian Foundation for Health and Rehabilitation

through the Norwegian Back Pain Association. The funders had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

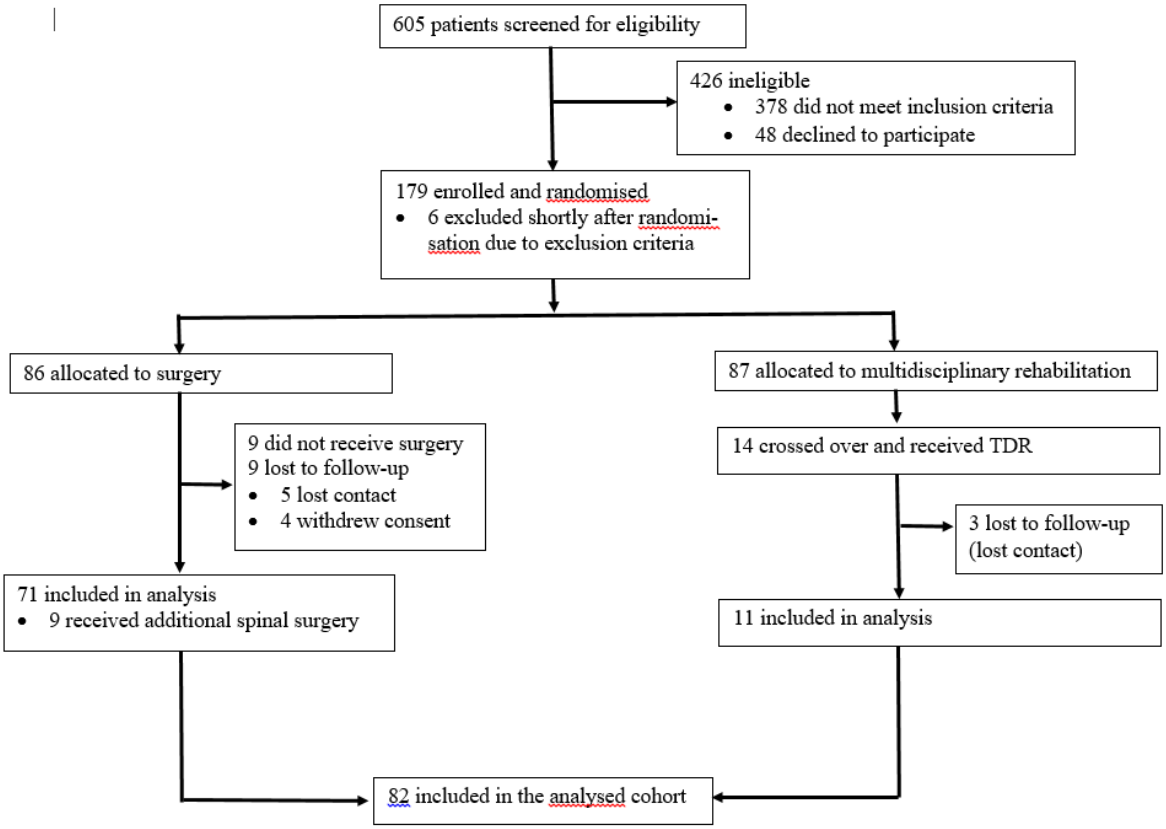
References

1. Jacobs W, Van der Gaag NA, Tuschel A, de Kleuver M, Peul W, Verbout AJ, Oner FC (2012) Total disc replacement for chronic back pain in the presence of disc degeneration. Cochrane database of systematic reviews (Online) 9:CD008326. doi: 10.1002/14651858.CD008326.pub2
2. Ding F, Jia Z, Zhao Z, Xie L, Gao X, Ma D, Liu M (2017) Total disc replacement versus fusion for lumbar degenerative disc disease: a systematic review of overlapping meta-analyses. Eur Spine J 26:806-815. doi: 10.1007/s00586-016-4714-y
3. Lemaire JP, Carrier H, Sariali el H, Skalli W, Lavaste F (2005) Clinical and radiological outcomes with the Charite artificial disc: a 10-year minimum follow-up. Journal of spinal disorders & techniques 18:353-359
4. David T (2007) Long-term results of one-level lumbar arthroplasty: minimum 10-year follow-up of the CHARITE artificial disc in 106 patients. Spine 32:661-666. doi: 10.1097/01.brs.0000257554.67505.45
5. Guyer RD, McAfee PC, Banco RJ, Bitan FD, Cappuccino A, Geisler FH, Hochschuler SH, Holt RT, Jenis LG, Majd ME, Regan JJ, Tromanhauser SG, Wong DC, Blumenthal SL (2009) Prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: five-year follow-up. The spine journal : official journal of the North American Spine Society 9:374-386. doi: 10.1016/j.spinee.2008.08.007
6. Skold C, Tropp H, Berg S (2013) Five-year follow-up of total disc replacement compared to fusion: a randomized controlled trial. Eur Spine J. doi: 10.1007/s00586-013-2926-y
7. Siepe CJ, Heider F, Wiechert K, Hitzl W, Ishak B, Mayer MH (2014) Mid- to long-term results of total lumbar disc replacement: a prospective analysis with 5- to 10-year follow-up. The spine journal : official journal of the North American Spine Society 14:1417-1431. doi: 10.1016/j.spinee.2013.08.028
8. Guyer RD, Pettine K, Roh JS, Dimmig TA, Coric D, McAfee PC, Ohnmeiss DD (2016) Five-Year Follow-Up of a Prospective, Randomized Trial Comparing Two Lumbar Total Disc Replacements. Spine 41:3-8. doi: 10.1097/brs.0000000000001168
9. Park SJ, Lee CS, Chung SS, Lee KH, Kim WS, Lee JY (2016) Long-Term Outcomes Following Lumbar Total Disc Replacement Using ProDisc-II: Average 10-Year Follow-Up at a Single Institute. Spine 41:971-977. doi: 10.1097/brs.0000000000001527
10. Furunes H, Storheim K, Brox JI, Johnsen LG, Skouen JS, Franssen E, Solberg TK, Sandvik L, Hellum C (2017) Total disc replacement versus multidisciplinary rehabilitation in patients with chronic low back pain and degenerative discs: Eight-year follow-up of a randomized controlled multicenter trial. The spine journal : official journal of the North American Spine Society. doi: 10.1016/j.spinee.2017.05.011
11. Hellum C, Johnsen LG, Gjertsen O, Berg L, Neckelmann G, Grundnes O, Rossvoll I, Skouen JS, Brox JI, Storheim K (2012) Predictors of outcome after surgery with disc prosthesis and rehabilitation in patients with chronic low back pain and degenerative disc: 2-year follow-up. Eur Spine J 21:681-690. doi: 10.1007/s00586-011-2145-3
12. Gornet MF, Schranck F, Wharton ND, Beall DP, Jones E, Myers ME, Hipp JA (2014) Optimizing success with lumbar disc arthroplasty. Eur Spine J 23:2127-2135. doi: 10.1007/s00586-014-3309-8
13. 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. doi: 10.1136/bmj.d2786
14. Fairbank JC, Pynsent PB (2000) The Oswestry Disability Index. Spine 25:2940-2952; discussion 2952

15. Statistics Norway (1984) Standard Classifications of Socioeconomic Status. In.
16. Statistics Norway (1998) Standard Classification of Occupations.
17. 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
18. 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
19. 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
20. Uden A, Astrom M, Bergenudd H (1988) Pain drawings in chronic back pain. Spine 13:389-392
21. Legaye J, Duval-Beaupere G, Hecquet J, Marty C (1998) Pelvic incidence: a fundamental pelvic parameter for three-dimensional regulation of spinal sagittal curves. Eur Spine J 7:99-103
22. 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
23. 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. doi: 10.1097/BRS.0b013e31817f19f7
24. 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
25. 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
26. 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. doi: 10.1259/0007-1285-65-773-361
27. Berg L, Neckelmann G, Gjertsen O, Hellum C, Johnsen LG, Eide GE, Espeland A (2012) Reliability of MRI findings in candidates for lumbar disc prosthesis. Neuroradiology 54:699-707. doi: 10.1007/s00234-011-0963-y
28. Geisser S (1993) Predictive Inference. Chapman and Hall, New York, NY
29. Laustsen AF, Bech-Azeddine R (2016) Do Modic changes have an impact on clinical outcome in lumbar spine surgery? A systematic literature review. Eur Spine J. doi: 10.1007/s00586-016-4609-y
30. Cougot B, Petit A, Paget C, Roedlich C, Fleury-Bahi G, Fouquet M, Menu P, Dubois C, Geraut C, Roquelaure Y, Tripodi D (2015) Chronic low back pain among French healthcare workers and prognostic factors of return to work (RTW): a non-randomized controlled trial. J Occup Med Toxicol 10:40. doi: 10.1186/s12995-015-0082-5
31. Grovle L, Haugen AJ, Keller A, Natvig B, Brox JI, Grotle M (2010) The bothersomeness of sciatica: patients' self-report of paresthesia, weakness and leg pain. Eur Spine J 19:263-269. doi: 10.1007/s00586-009-1042-5
32. McGirt MJ, Sivaganesan A, Asher AL, Devin CJ (2015) Prediction model for outcome after low-back surgery: individualized likelihood of complication, hospital readmission, return to work, and 12-month improvement in functional disability. Neurosurg Focus 39:E13. doi: 10.3171/2015.8.focus15338
33. Nordeman L, Gunnarsson R, Mannerkorpi K (2014) Prognostic factors for work ability in women with chronic low back pain consulting primary health care: a 2-year prospective longitudinal cohort study. Clin J Pain 30:391-398. doi: 10.1097/AJP.0b013e3182a0dd06
34. Johnsen LG, Hellum C, Nygaard OP, Storheim K, Brox JI, Rossvoll I, Leivseth G, Grotle M (2013) Comparison of the SF6D, the EQ5D, and the oswestry disability index in patients with chronic low back pain and degenerative disc disease. BMC Musculoskelet Disord 14:148. doi: 10.1186/1471-2474-14-148
35. van Hooff ML, Mannion AF, Staub LP, Ostelo RW, Fairbank JC (2016) Determination of the Oswestry Disability Index score equivalent to a "satisfactory symptom state" in patients undergoing surgery for degenerative disorders of the lumbar spine-a Spine Tango registry-based study. The spine

journal : official journal of the North American Spine Society 16:1221-1230. doi: 10.1016/j.spinee.2016.06.010
36. Peng B, Fu X, Pang X, Li D, Liu W, Gao C, Yang H (2012) Prospective clinical study on natural history of discogenic low back pain at 4 years of follow-up. Pain physician 15:525-532
37. Geisler FH, Guyer RD, Blumenthal SL, McAfee PC, Cappuccino A, Bitan F, Regan JJ (2008) Patient selection for lumbar arthroplasty and arthrodesis: the effect of revision surgery in a controlled, multicenter, randomized study. J Neurosurg Spine 8:13-16. doi: 10.3171/spi-08/01/013

Figure 1. Trial profile



1

2

Figure 2. Prediction matrix

		Low education		High education	
		Comorbidity	No comorbidity	Comorbidity	No comorbidity
≥ 12 months sick leave	ODI ≥ 50	1 % (0-4)	9 % (3-15)	4 % (0-8)	24 % (15-33)
	ODI < 50	4 % (0-8)	25 % (16-35)	12 % (5-19)	52 % (41-63)
< 12 months sick leave	ODI ≥ 50	7 % (2-13)	38 % (28-49)	20 % (12-29)	67 % (56-77)
	ODI < 50	22 % (13-31)	68 % (58-78)	47 % (36-58)	87 % (80-94)

Probability of working (95 % CI) at long-term follow-up after total disc replacement using a probability matrix model. Educational level (≤ 9 years or > 9 years, presence of comorbidity, duration of sick leave before treatment (< 12 months or ≥ 12 months) and Oswestry Disability Index (ODI, < 50 points or ≥ 50 points).

Table 1. Baseline characteristics of analysed patient cohort

	n	%
Socioeconomic variables		
Manual worker (yes/no) (n=76)	31/45	41
Educational level (n=82)		
Primary and secondary school (9 years)	18	22
High school (12 years)	44	54
University/college	20	24
Working (yes/no) (n=82)	26/56	32
Duration of sick leave (months) (median, range) (n=79)	12 (0-70)	
Current smoker (yes/no) (n=81)	38/43	47
Gender (female/male) (n=82)	40/42	49
Age (median, range) (n=82)	41 (25-54)	
Clinical variables		
Prior surgery (yes/no) (n=82)	26/56	32
Affected level (n=82)		
L4/L5	17	21
L5/S1	39	48
L4/L5 and L5/S1	26	32
Comorbidity (yes/no) (n=82)	20/62	24
Oswestry Disability Index (median, range) (n=82)	40.0 (28.0-70.0)	
Body Mass Index (median, range) (n=80)	25.1 (18.5-35.4)	
Psychological variables and pain		
Hopkins Symptoms Checklist - 25 (median, range) (n=77)	1.68 (1.00-3.12)	
Fear Avoidance Beliefs Questionnaire - work (median, range) (n=74)	29.0 (2.0-42.0)	
Fear Avoidance Beliefs Questionnaire - physical (median, range) (n=75)	14.0 (2.0-24.0)	
Short Form - 36 Mental Component Summary (median, range) (n=77)	49.6 (13.0-71.4)	
Back pain (Visual Analogue Scale) (median, range) (n=80)	67.5 (19.0-97.0)	
Pain drawing (below waist/above waist) (n=77)	61/16	79
Duration of back pain (years) (median, range) (n=71)	4.0 (0.2-25.0)	
Daily consumption of narcotics (yes/no) (n=61)	25/36	41
Radiological variables		
Pelvic incidence (median, range) (n=74)	50.0 (25.0-79.0)	
Modic changes (n=81)		
Not present	13	16
Type 1	23	28
Type 2	30	37
Types 1 and 2	15	19
> 50 of vertebral body height (yes/no) (n=81)	27/54	33
Disc height reduction > 40 % (yes/no) (n=81)	55/26	68
Nucleus pulposus grade 3 or 4 (yes/no) (n=81)*	72/9	89
Facet arthropathy grade 2 or 3 (yes/no) (n=81)**	9/72	11
Posterior high intensity zone (yes/no) (n=81)	43/38	53

* Luoma et al. (ref)

** Fujiwara et al. (ref)

Table 2. Association between baseline characteristics and a clinically important improvement of 15 ODI points at long-term follow-up of patients undergoing TDR (achieved by 52 of 82 patients (63 %)).

	Univariate logistic regression		
	OR	*95 % CI*	*P*
Socioeconomic variables			
Manual worker (no/yes)	1.65	0.64-4.22	0.30
Educational level			
Higher education (> 9 years vs ≤ 9 years)	1.13	0.39-3.33	0.82
Working (yes/no)	1.13	0.43-3.00	0.80
Duration of sick leave			
< 12 months (vs ≥ 12 months)	1.58	0.62-4.02	0.34
Current smoker (yes/no)	1.26	0.51-3.12	0.62
Gender (female)	1.41	0.57-3.49	0.45
Age ≥ 40 (yes/no)	1.48	0.60-3.65	0.40
Clinical variables			
Prior surgery (no/yes)	1.81	0.70-4.70	0.22
Affected level			
L4/L5 (vs L4/L5 and L5/S1)	1.07	0.28-4.05	0.93
L5/S1 (vs L4/L5 and L5/S1)	0.58	0.20-1.64	0.30
L4/L5 and L5/S1			
Comorbidity (yes/no)	1.10	0.38-3.14	0.87
Oswestry Disability Index ≥ 50 (yes/no)	2.70	0.70-10.48	0.15
Body Mass Index < 25 (yes/no)	2.03	0.80-5.13	0.14
Psychological variables and pain			
Hopkins Symptoms Checklist - 25	0.89	0.35-2.24	0.80
Fear Avoidance Beliefs Questionnaire - work	0.99	0.95-1.04	0.80
Fear Avoidance Beliefs Questionnaire - physical	0.95	0.87-1.05	0.32
Short Form - 36	0.99	0.96-1.03	0.60
Back pain (Visual Analogue Scale)	1.01	0.98-1.04	0.76
Pain drawing (above waist/below waist)	2.81	0.72-10.91	0.14
Duration of back pain (≥ 5 years/< 5 years)	1.98	0.69-5.65	0.20
Daily consumption of narcotics (yes/no)	1.00	0.34-2.88	0.99
Radiological variables			
Pelvic incidence > 55 (yes/no)	1.23	0.44-3.41	0.70
Modic changes			
Present (vs not present)	5.04	1.39-18.21	**0.01**
Type 1 (vs not type 1)	1.56	0.63-3.89	0.34
Type 2 (vs not type 2)	1.77	0.71-4.41	0.22
Types 1 and 2 (vs not types 1 and 2)	1.22	0.37-3.98	0.74
> 50 of craniocaudal diameter (yes/no)	3.79	1.25-11.49	**0.02**
Disc height reduction > 40 % (yes/no)	1.39	0.53-3.62	0.50
Nucleus pulposus grade 3 or 4 (no/yes) *	1.20	0.28-5.20	0.81
Facet arthropathy grade 2 or 3 (yes/no) **	1.20	0.28-5.20	0.81
Posterior high intensity zone (no/yes)	1.26	0.51-3.12	0.62

* Luoma et al. (ref)

** Fujiwara et al. (ref)

Table 3. Association between baseline characteristics and employment at long-term follow-up of patients undergoing TDR (42 of 82 patients (51 %) were employed at follow-up).

	Univariate logistic regression			Multiple logistic regression		
	OR	*95 % CI*	*P*	*OR*	*95 % CI*	*P*
Socioeconomic variables						
Manual worker (no/yes)	1.3	0.5-3.3	0.54			
Educational level						
Higher education (> 9 years vs ≤ 9 years)	3.6	1.1-11.2	0.03	3.2	0.8-12.1	0.84
Working (yes/no)	2.3	0.9-6.2	0.08			
Duration of sick leave						
< 12 months (vs ≥ 12 months)	4.1	1.6-10.6	0.003	6.3	2.0-19.6	0.002
Gender (male)	1.3	0.6-3.2	0.51			
Current smoker (no/yes)	1.6	0.6-3.8	0.32			
Age < 40 (yes/no)	2.0	0.8-5.0	0.12			
Clinical variables						
Prior surgery (no/yes)	2.1	0.8-5.5	0.12			
Affected level						
L4/L5 (vs L4/L5 and L5/S1)	1.8	0.6-4.8	0.27			
L5/S1 (vs L4/L5 and L5/S1)	1.5	0.4-5.2	0.50			
L4/L5 and L5/S1						
Comorbidity (no/yes)	4.4	1.4-13.8	0.01	7.7	2.0-30.5	0.003
Oswestry Disability Index ≥ 50 (no/yes)	3.6	1.0-12.5	0.04	3.4	0.8-15.2	0.11
Body Mass Index ≥ 25 (no/yes)	1.2	0.5-2.9	0.65			
Psychological variables and pain						
Hopkins Symptoms Checklist - 25	1.0	0.4-2.4	0.95			
Fear Avoidance Beliefs Questionnaire - work	0.9	0.9-1.0	0.01			
Fear Avoidance Beliefs Questionnaire - physical	0.9	0.9-1.0	0.16			
Short Form - 36	1.0	1.0-1.0	0.80			
Back pain (Visual Analogue Scale)	1.0	1.0-1.0	0.86			
Pain drawing (above (and below) waist/below waist)	1.2	0.4-3.8	0.70			
Duration of back pain (≥ 5 years/< 5 years)	1.4	0.5-3.6	0.52			
Daily consumption of narcotics (no/yes)	1.4	0.5-4.0	0.50			
Radiological variables						
Pelvic incidence ≥ 55 (yes/no)	2.2	0.8-6.0	0.14			
Modic changes						
Present (vs not present)	1.3	0.4-4.3	0.65			
Type 1 (vs not type 1)	1.3	0.5-3.1	0.56			
Type 2 (vs not type 2)	0.8	0.3-1.8	0.55			
Types 1 and 2 (vs not types 1 and 2)	0.8	0.3-2.4	0.66			
> 50 of vertebral body height (yes/no)	2.0	0.8-5.1	0.16			
Disc height reduction < 40 % (yes/no)	1.1	0.4-2.9	0.81			
Nucleus pulposus grade 3 or 4 (yes/no) *	1.4	0.4-5.6	0.64			
Facet arthropathy grade 2 or 3 (yes/no) **	1.2	0.3-4.8	0.81			
Posterior high intensity zone (no/yes)	1.3	0.5-3.1	0.56			

* Luoma et al. (ref)

** Fujiwara et al. (ref)