

# Surgical Treatment of Degenerative Disk Disease in Three Scandinavian Countries: An International Register Study Based on Three Merged National Spine Registers

Mikkel Østerheden Andersen, MD<sup>1,2</sup>, Peter Fritzell, MD, PhD<sup>3,4,5</sup>, Søren Peter Eiskjaer, MD<sup>6</sup>, Tobias Lagerbäck, MD<sup>7</sup>, Olle Hägg, MD, PhD<sup>8</sup>, Dennis Nordvall, MSc<sup>5</sup>, Greger Lönne, MD<sup>9,10</sup>, Tore Solberg, MD, PhD<sup>11,12</sup>, Wilco Jacobs, MSc, PhD<sup>13</sup>, Miranda van Hooff, MSc, PhD<sup>14,15</sup>, Paul Gerdhem, MD, PhD<sup>7</sup>, and Martin Gehrchen, MD, PhD<sup>16</sup>

Global Spine Journal  
2019, Vol. 9(8) 850-858  
© The Author(s) 2019  
Article reuse guidelines:  
sagepub.com/journals-permissions  
DOI: 10.1177/2192568219838535  
journals.sagepub.com/home/gsj



## Abstract

**Study Design:** Observational study of prospectively collected data.

**Objectives:** Patients with chronic low back pain resistant to nonoperative treatment often face a poor prognosis for recovery. The aim of the current study was to compare the variation and outcome of surgical treatment of degenerative disc disease in the Scandinavian countries based on The International Consortium for Health Outcomes Measurement core spine data sets.

**Methods:** Anonymized individual level data from 3 national registers were pooled into 1 database. At the time of surgery, the patient reports data on demographics, lifestyle topics, comorbidity, and data on health-related quality of life such as Oswestry Disability Index, Euro-Qol-5D, and back and leg pain scores. The surgeon records diagnosis, type of surgery performed, and complications. One-year follow-ups are obtained with questionnaires. Baseline and 1-year follow-up data were analyzed to expose any differences between the countries.

**Results:** A total of 1893 patients were included. At 1-year follow-up, 1315 (72%) patients responded. There were statistically significant baseline differences in age, smoking, comorbidity, frequency of previous surgery and intensity of back and leg pain. Isolated fusion was the primary procedure in all the countries ranging from 84% in Denmark to 76% in Sweden. There was clinically relevant improvement in all outcome measures except leg pain.

**Conclusions:** In homogenous populations with similar health care systems the treatment traditions can vary considerably. Despite variations in preoperative variables, patient reported outcomes improve significantly and clinically relevant with surgical treatment.

## Keywords

degenerative disc disease, degenerative, fusion, lumbar interbody fusion, disc replacement

<sup>1</sup> Center for Spine Surgery & Research, Region of Southern Denmark, Middelfart, Denmark

<sup>2</sup> University of Southern Denmark, Odense, Denmark

<sup>3</sup> Capio St Göran Hospital, Stockholm, Sweden

<sup>4</sup> Uppsala University, Uppsala, Sweden

<sup>5</sup> Qulturum Center for Learning and Innovation in Healthcare, Jönköping, Sweden

<sup>6</sup> Aalborg University Hospital, Aalborg C, Denmark

<sup>7</sup> Karolinska University Hospital and Karolinska Institutet, Huddinge, Sweden

<sup>8</sup> Spine Center Göteborg, Göteborg, Sweden

<sup>9</sup> Innlandet Hospital Trust, Lillehammer, Norway

<sup>10</sup> St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway

<sup>11</sup> University Hospital of Northern Norway, Tromsø, Norway

<sup>12</sup> The Arctic University of Norway, Tromsø, Norway

<sup>13</sup> The Health Scientist, The Hague, the Netherlands

<sup>14</sup> Sint Maartenskliniek, Nijmegen, the Netherlands

<sup>15</sup> Radboud University Medical Center, Nijmegen, the Netherlands

<sup>16</sup> Rigshospitalet University of Copenhagen, Copenhagen, Denmark

## Corresponding Author:

Martin Gehrchen, Spine Unit, Department of Orthopedic Surgery, Rigshospitalet, Copenhagen University Hospital, Blegdamsvej 9, 2100 Copenhagen, Denmark.

Email: pedicle@mac.com



## Introduction

Many patients with chronic intractable lower back pain of discogenic origin do not recover with conservative, nonoperative management alone. Consequently, this patient group is confronted with the option of living with persistent back pain or undergoing surgical spinal fusion or total disc replacement. However, the success of surgical treatment versus usual nonoperative management is debatable.<sup>1</sup>

In the recent published guidelines from the British National Institute for Health and Care Excellence (NICE) it is recommended to “not offer spinal fusion for people with low back pain unless as part of a randomized controlled trial” despite the fact that they identified studies indicating that spinal fusion was more beneficial for some elements of pain, function, and quality of life (QoL) and that health care use was lower. NICE describes the evidence based on these studies as weak due to low numbers of patients, large crossover, and in-case selection bias.<sup>1</sup>

Clinical registries collect data from everyday practice and can evaluate different treatment strategies by linking practice-based variation to treatment effectiveness. These registries thus increase the external validity. Studies based on such data allow surgeons and patients to choose type of surgery according to their preferences.<sup>2</sup>

The International Consortium for Health Outcomes Measurement (ICHOM) cooperation aims at defining core data sets in different diagnostic entities to enable relevant comparisons of outcome between clinics and countries.<sup>3</sup> The spine surgery registries of Denmark, Norway, and Sweden were among the collaborators in this effort and use similar sets of patient-reported outcome measures (PROMs). The Scandinavian population is genetically similar, and the countries have similar social security systems, similar language, public-based health care and health insurance systems, facilitating comparative studies.<sup>4</sup>

The aims of this study were (1) to compare variation in surgical treatment of lumbar degenerative disk disease (DDD) in terms of surgical selection criteria (preoperative patient characteristics), (2) to assess if practice-based variations were associated to different patient-reported outcomes in a large combined registry cohort from the 3 Scandinavian countries, and (3) to analyze the data with regard to factors influencing the result of surgical treatment.

## Methods

This is an observational study, reviewing prospectively collected data from the national spine registries of Denmark (DaneSpine), Norway (NORspine), and Sweden (Swespine). Inclusion criteria were: age between 18 and 65 years, either had fusion surgery or disc replacement and operated for lumbar DDD between January 2011 and December 2013.

The diagnosis of lumbar DDD was based on the surgeons' clinical judgment, x-ray, and magnetic resonance imaging (MRI).

This study was approved by ethical review boards in Denmark (Projekt-ID: S-20 160 091), Norway (REC South-east B: 2014/2219), and Sweden (Dnr 2015/181-31). The study was conducted and reported in accordance with the study protocol, which is available at [clinicaltrials.gov](http://clinicaltrials.gov) (ID: NCT02980822).

## The Registries

All 3 national spine registries are designed for quality control and research. The participation is voluntary for the surgical departments as well as the patient. At the time of admission for surgery (baseline), the patient self-reports data on demographics, life style matters, comorbidity, and PROMs with the use of questionnaires. During the hospital stay, the surgeon records diagnosis, type of surgery performed, and perioperative complications. One-year follow-up does not involve any health professionals at the treating hospital. Questionnaires are distributed, completed at home by the patients and returned in prestamped envelopes. The oldest registry is Swespine, which has included individuals treated with lumbar surgery since 1998. Swespine covers approximately 90% of the surgical units in Sweden. Completeness, the proportion of operated patients reported to Swespine, was approximately 75% in the study period.<sup>5</sup> NORspine, is based on the concept of the Swespine register and was founded in 2007. Coverage in NORspine is approximately 95%. The completeness is approximately 65% in the study period.<sup>6</sup> DaneSpine was acquired by the Danish Spine Society from the Swedish Society of Spinal Surgeons in 2009 and has successively been implemented. Coverage is approximately 80%. The completeness is approximately 62% in the study period.<sup>7,8</sup>

## Patient-Reported Outcome Measures

The primary outcome measure was the Oswestry Disability Index (ODI, version 2.1) ranging from 0 (no disability) to 100 (bedridden).<sup>9</sup> The ODI is a standard for measuring pain-related disability in persons with low back pain.

Secondary outcome measures were numeric rating scales (NRS) for back and leg pain, ranging from 0 (no pain) to 10 (worst conceivable pain).<sup>10</sup> Health-related quality of life was measured with the 3 level Euro-Qol-5D (EQ-5D-3L) ranging from -0.596 to 1, with higher scores indicating better quality of life (according to the British tariff—UK-Time Trade-Off).<sup>11</sup>

NORspine used the NRS for leg and back pain,<sup>12</sup> while Swespine and DaneSpine used the visual analogue scale (VAS) for back and leg pain, ranging from 0 to 100.<sup>13</sup> Conversion to the NRS was done by dividing the VAS score by 10 with stochastic approximation of decimals to the closest integer.

## Data Handling and Analysis

Anonymized individual level data from all 3 registers were pooled into 1 database, and the cohort was divided by country for comparison. Missing or out-of-range data on gender, age,

height, or weight were deleted. Cases with missing date of surgery and follow-up were excluded.

**Nonresponse Analysis.** A nonresponse analysis was performed by comparing all available baseline variables between those who responded to the 1-year follow-up to those who did not.

### Statistics

Analysis of baseline data included PROM-scores, age at date of surgery, sex, height, weight, smoking habits, sick leave, and duration of leg and back pain presented as mean (with SD or 95% confidence interval), or proportions. Variables were analyzed by analysis of variance (ANOVA), chi-square, or logistic regression tests. Data is presented as crude (unadjusted) to elucidate any differences between the countries, and adjusted for case mix (baseline data) with linear regression analysis.

Comparisons of the mean change of the PROMs at 1 year were analyzed by ANOVA. The minimal clinically important difference (MCID) between groups was defined as 15 for ODI and 2.0 for NRS back pain and leg pain.<sup>14,15</sup>

Logistic regression analysis was used for predictive modeling. The difference in the EQ-5D-3L score from before surgery to 1 year after was chosen as the dependent variable with cut point 0.19 (1/2\*SD).<sup>16,17</sup> The following variables were entered as independent variables, age, gender, smoking, body mass index, country (2 dummy variables), duration of back pain before surgery, duration of leg pain before surgery, number of previous spine surgeries, EQ-5D-3L score before surgery dichotomized with 0.4 as cut point, ODI score before surgery, on sick leave or not before surgery, on pain medication or not before surgery, of back pain, duration of leg pain, number of previous spine surgeries, type of surgery, EQ5D-3L anxiety score before surgery, number of levels included in the fusion. For the logistic regression analysis, a complete data set were available for 1184 patients. Multicollinearity between the independent variables was investigated using the variance inflation factor (VIF; a value >4 was considered index of multicollinearity). The Hosmer-Lemeshow test for goodness of fit of the logistic regression model, the Akaike information criteria (AIC), receiver operating characteristic (ROC) curves and the area under the curve (AUC) were used to evaluate the performance of the logistic regression models.<sup>18</sup> Stepwise regression (combined forward and backward, significance level 0.05) was used to construct the final model. A Shiny app (<https://cran.r-project.org/package=shiny>) was programmed using the data from the final logistic regression model. The dynamic nomogram is accessible at (<https://dynamisknomogramse.shinyapps.io/dynnomapp/>).

### Results

A total of 1893 patients were included. From Denmark, 392 patients were included and correspondingly from Norway 300 and Sweden 1147. At 1-year follow-up, 1315 (72%) responded (Denmark, n = 259 (66%); Norway, n = 164 (55%); and

Sweden, n = 892 (78%). Figure 1 shows the selection process. There were statistically significant differences between the countries in several baseline variables (Table 1). The Danish patients were older than their Scandinavian peers. Fewer were smoking in Sweden. In all 3 countries, more than 80% of the patients had more than 1-year duration of preoperative back pain. Health-related quality of life (EQ-5D-3L  $\pm$  SD) was better in Norway (0.36  $\pm$  0.30) and Sweden (0.32  $\pm$  0.33) compared with Denmark (0.29  $\pm$  0.33), and NRS leg pain and back pain intensity were significantly higher in Denmark as was the preoperative comorbidity. There was no significant difference in preoperative ODI between the countries (Table 1). The frequency of previous surgery varied significantly between the countries ranging from 45% in Denmark to 37% in Sweden (Table 2).

### Type of Surgery

There was a significant variation in the surgical technique between the 3 countries. Isolated fusion was the primary procedure in all the countries: Denmark (84%), Norway (82%), and Sweden (76%). Total disc replacement (TDR) was more frequent in Sweden (24%) versus Norway (18%) and Denmark (16%) (Table 2). Comparing patients treated with fusion to TDR the rates of previous surgery and the pain-related disability were higher in the fused group and health-related quality of life were lower (Table 2). The ODI and EQ-5D-3L scores improved more in favor of TDR than fusion surgery, with a mean difference of ODI score of 6 and 0.09 in EQ-5D-3L, but with differences between the countries (Table 2).

### Outcome at 1 Year

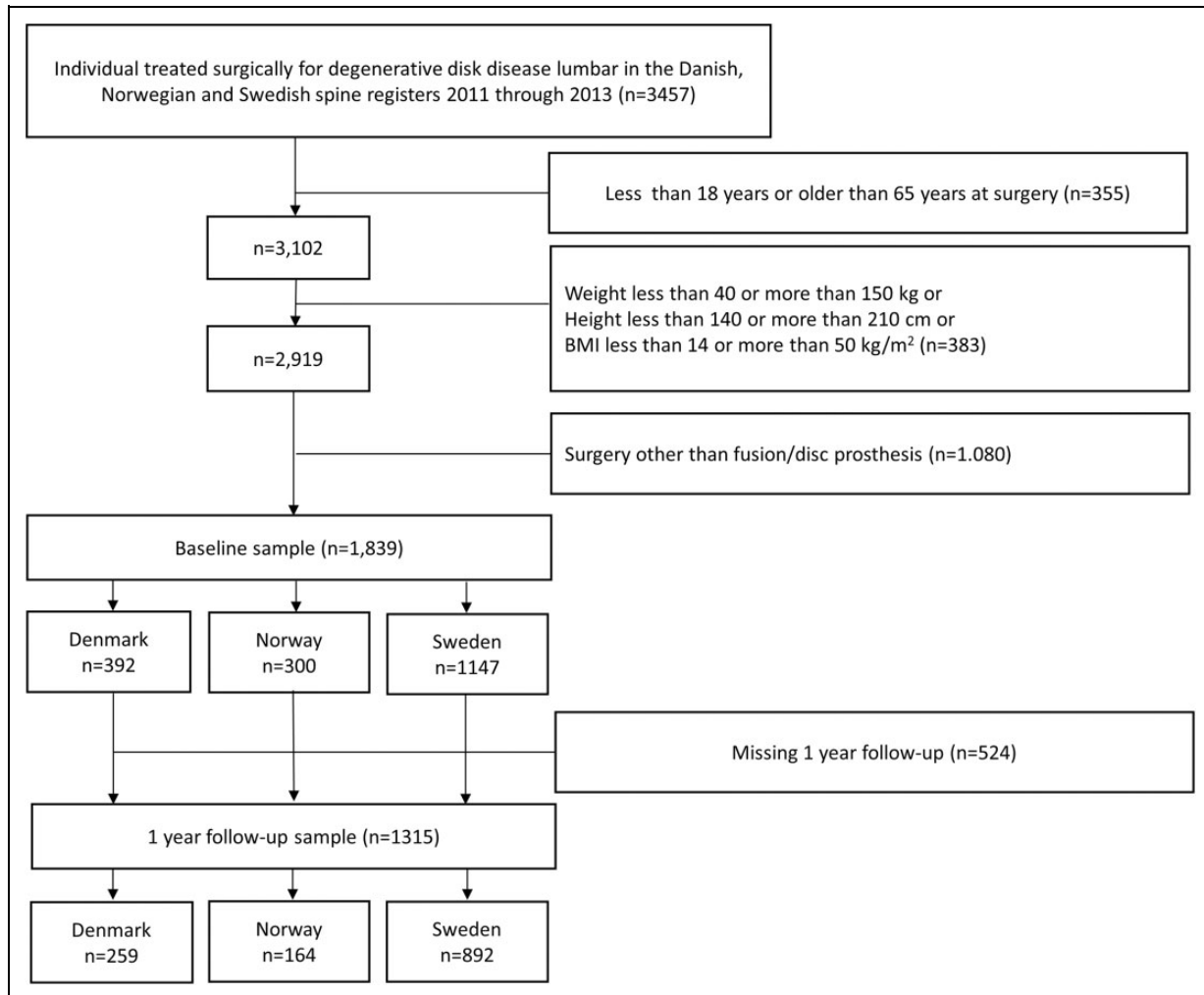
There were overall statistically significant mean improvements in all outcome measures (Tables 2 and 3). Apart from improvement in leg pain, all outcome measures reached clinically relevant changes. Between the countries, there were significant differences in all outcomes except EQ-5D-3L and EQ-5D-3L improvement. After case-mix adjustment, all outcomes had significant differences (Table 3). Norway had less mean improvement in all outcomes when compared with Denmark and Sweden.

### Nonresponders

The nonresponders represented 28% of the cohort. They were 2 years younger than the responders and had a higher proportion of males and smokers (Table 4).

### Predictive Modeling

All VIFs were <4. The following variables had a significant predictive value in the logistic regression analysis: age, duration of back pain before surgery, number of previous spine surgeries, EQ5D score before surgery, ODI score before surgery, on sick leave before surgery, on pain medication before surgery, type of surgery (fusion or disc prosthesis). Odds ratios,



**Figure 1.** Flowchart of selection process.

**Table 1.** Baseline Characteristics Presented as Mean (SD) or Proportions.

Characteristic	Total	Denmark	Norway	Sweden	P
Number of patients, n	1839	392	300	1147	
Age, years, mean (SD)	45.3 (9.9)	48.1 (10.7)	43.9 (10.3)	44.7 (9.3)	<.001 <sup>a</sup>
BMI, kg/m <sup>2</sup> , mean (SD)	26.2 (3.9)	26.6 (4.2)	26.4 (4.3)	26.0 (3.7)	<.02 <sup>a</sup>
Females, n (%)	1039 (56)	238 (61)	165 (55)	636 (55)	.16 <sup>b</sup>
Smokers, n (%)	282 (15)	107 (27)	81 (27)	94 (8)	<.001 <sup>b</sup>
Preoperative pain medication, n (%)	1635 (90)	326 (84)	270 (93)	1039 (91)	<.001 <sup>b</sup>
Neurological comorbidity, n (%)	20 (1)	3 (1)	4 (1)	13 (1%)	.75 <sup>b</sup>
Heart comorbidity, n (%)	26 (1.4)	14 (3.6)	2 (0.7)	10 (0.9)	<.001 <sup>b</sup>
Cancer comorbidity, n (%)	9 (0.5)	6 (1.5)	1 (0.3)	2 (0.2)	.004 <sup>b</sup>
Preoperative duration of leg pain >12 months, n (%)	1136 (63)	253 (68)	176 (61)	707 (62)	.12 <sup>b</sup>
Preoperative duration of back pain >12 months, n (%)	1613 (88)	331 (85)	256 (88)	1026 (90)	.04 <sup>b</sup>
ODI, mean (SD)	43 (14.2)	44 (14.7)	42 (14.0)	44 (14.1)	.26 <sup>a</sup>
NRS leg pain, mean (SD)	4.3 (3.0)	5.7 (2.8)	4.2 (3.2)	3.9 (2.9)	<.001 <sup>a</sup>
NRS back pain, mean (SD)	6.3 (2.2)	6.9 (1.8)	5.3 (2.9)	6.4 (2.0)	<.001 <sup>a</sup>
EQ-5D-3L, mean (SD)	0.32 (0.32)	0.29 (0.33)	0.36 (0.30)	0.32 (0.33)	.02 <sup>a</sup>
Responding at the 1-year follow-up, n (%)	1315 (72)	259 (66)	164 (55)	892 (78)	<.001 <sup>b</sup>

Abbreviations: SD, standard deviation; BMI, body mass index; ODI, Oswestry Disability Index; NRS, numeric rating scale.

<sup>a</sup>Analysis of variance *F* test.

<sup>b</sup>Pearson's chi-square test.

**Table 2.** Baseline Characteristics for Patients Treated With Either Spinal Fusion or Disc Replacement Presented as Mean (SD) or Proportions.

Characteristic	Total	Denmark	Norway	Sweden	P
Fusion surgery, n (%)	1444 (79)	329 (84)	245 (82)	870 (76)	<.001 <sup>a</sup>
Frequency of prior surgery, %	39	45	40	37	<.001 <sup>a</sup>
Pre-ODI, mean (SD)	44 (14.3)	44 (15.0)	42 (13.6)	45 (14.2)	.02 <sup>b</sup>
Post-ODI, mean (SD)	26 (20.2)	28 (20.6)	29 (18.9)	26 (20.3)	.12 <sup>b</sup>
Δ ODI	17	16	13	19	
Pre-EQ-5D-3L, mean (SD)	0.31 (0.32)	0.28 (0.33)	0.36 (0.30)	0.30 (0.33)	.006 <sup>b</sup>
Post-EQ-5D-3L, mean (SD)	0.60 (0.36)	0.56 (0.36)	0.61 (0.31)	0.61 (0.36)	.32 <sup>b</sup>
Δ EQ-5D-3L	0.29	0.28	0.25	0.31	
Disc replacement, n (%)	395 (21)	63 (16)	55 (18)	277 (24)	<.001 <sup>a</sup>
Frequency of prior surgery, %	17	21	18	16	.25 <sup>a</sup>
Pre-ODI, mean (SD)	40 (13.4)	42 (13.5)	42 (15.9)	39 (12.9)	.18 <sup>b</sup>
Post-ODI, mean (SD)	17 (17.2)	15 (17.1)	30 (16.8)	17 (16.8)	.001 <sup>b</sup>
Δ ODI	23	27	12	22	
Pre-EQ-5D-3L, mean (SD)	0.36 (0.32)	0.32 (0.30)	0.31 (0.31)	0.37 (0.33)	.23 <sup>b</sup>
Post-EQ-5D-3L, mean (SD)	0.74 (0.28)	0.75 (0.29)	0.63 (0.32)	0.74 (0.28)	.17 <sup>b</sup>
Δ EQ-5D-3L	0.38	0.43	0.32	0.37	

Abbreviations: SD, standard deviation; ODI, Oswestry Disability Index; Δ, change in score.

<sup>a</sup>Pearson's chi-square test.

<sup>b</sup>Analysis of variance F test.

**Table 3.** Postoperative Outcome and Change in Outcome From Baseline to 1 Year Postoperatively Shown as Mean (SD).<sup>a</sup>

	Total	Denmark	Norway	Sweden	P <sup>b</sup>	P <sup>c</sup>
Number of patients	1315	259	164	892		
ODI	24.6 (20.0)	25.7 (20.5)	29.1 (18.5)	23.4 (19.9)	<.002	<.001
Δ ODI	-18.3 (17.9)	-18.3 (17.9)	-12.1 (16.5)	-19.4(17.7)	<.001	<.001
NRS leg pain	2.4 (2.9)	3.1 (3.0)	2.7 (3.0)	2.1 (2.8)	<.001	<.001
Δ NRS leg pain	1.8 (3.3)	2.6 (3.3)	1.2 (3.8)	1.6 (3.1)	<.001	<.001
NRS back pain	3.2 (2.9)	4.0 (2.9)	3.6 (3.0)	2.9 (2.9)	<.001	<.001
Δ NRS back pain	3.1 (3.1)	2.9 (2.9)	1.7 (3.2)	3.4 (3.1)	<.001	<.001
EQ-5D-3L	0.63 (0.35)	0.59 (0.36)	0.61 (0.31)	0.64 (0.35)	.13	<.001
Δ EQ-5D-3L	0.30 (0.38)	0.30 (0.37)	0.24 (0.35)	0.30 (0.38)	.12	<.001

Abbreviations: SD, standard deviation; ODI, Oswestry Disability Index; Δ, change in score.

<sup>a</sup>P values calculated using analysis of covariance.

<sup>b</sup>Nonadjusted P value.

<sup>c</sup>Adjusted for age, body mass index, smoking, duration of back pain, and preoperative value of the dependent variable.

confidence intervals, coding, and increments are listed in Table 5. The P value for the Hosmer-Lemeshow test was 0.29. AIC was 1394 for the final model. AUC was 0.76 (95% CI 0.73-0.79) corresponding to an acceptable discrimination ability (see Figure 2).

## Discussion

This study represents, to our knowledge, the worlds' largest observational study with n = 1839 patients operated for DDD, reporting outcomes based on the ICHOM-recommended value set.

In a group of selected patients with chronic lower back pain where nonoperative treatment has failed, it is encouraging that the patients improve significantly and clinically relevant. The change in favor of TDR is in line with the study published by Berg et al<sup>19</sup> who reported that TDR was superior to spinal fusion in clinical outcome.

Even though the Scandinavian countries have almost similar public health care systems the selection criteria for surgery due to discogenic pain in terms of demographic characteristics, pain intensity, and disability were dissimilar. Furthermore, we found a significant practice variation, that is, the use of disc replacement surgery was significantly higher in Sweden compared with Denmark and Norway. This demonstrates that even in homogenous populations with similar health care systems the treatment traditions can vary considerably.

Fusion as a treatment option for patients with chronic low back pain is still a controversial topic. Unfortunately, there are only very few randomized studies comparing surgical versus nonsurgical treatment. In 2000, Möller and Hedlund<sup>20</sup> published a trial of 77 patients randomized to surgery or to an exercise program. The patients allocated to surgery reported greater benefits at 2 years in terms of ODI scores. In a Swedish study, Fritzell et al<sup>21</sup> randomized 222 patients to different surgical groups of equal size and 72 patients to physiotherapy. In

**Table 4.** Baseline Characteristics of Responders and Nonresponders Presented as Mean (SD) or Proportions.<sup>a</sup>

	Total	Denmark	Norway	Sweden
Age, years (n) responders	45.8 (1315)	49.1 (259)	44.8 (164)	45.1 (892)
Age, years (n) nonresponders	43.9 (524)	46.4 (133)	42.7 (136)	43.3 (255)
<i>p</i> <sup>b</sup>	<.001	.02	.082	.006
Females, n (%) responders	762 (58)	159 (61)	94 (57)	509 (57)
Females, (n) (%) nonresponders	277 (53)	70 (59)	71 (52)	127 (50)
<i>p</i> <sup>c</sup>	.05	.70	.38	.04
Smoking, n (%) responders	1302 (14)	259 (27)	163 (25)	880 (7)
Smoking, n (%) nonresponders	518 (20)	132 (28)	133 (30)	253 (11)
<i>p</i> <sup>c</sup>	<.001	.83	.35	.04
ODI (n) responders	43 (1301)	44 (259)	41 (162)	43 (880)
ODI (n) nonresponders	45 (507)	43 (132)	43 (162)	46 (253)
<i>p</i> <sup>b</sup>	.01	.86	.23	<.001
EQ-5D-3L (n) responders	0.33 (1304)	0.30 (252)	0.37 (164)	0.34 (888)
EQ-5D-3L (n) nonresponders	0.28 (502)	0.26 (125)	0.34 (125)	0.26 (252)
<i>p</i> <sup>a</sup>	.002	.35	.39	<.001

Abbreviations: SD, standard deviation; ODI, Oswestry Disability Index.

<sup>a</sup>*P* values represent comparisons of responders and non-responders.

<sup>b</sup>Analysis of variance *F* test.

<sup>c</sup>Pearson's chi-square test.

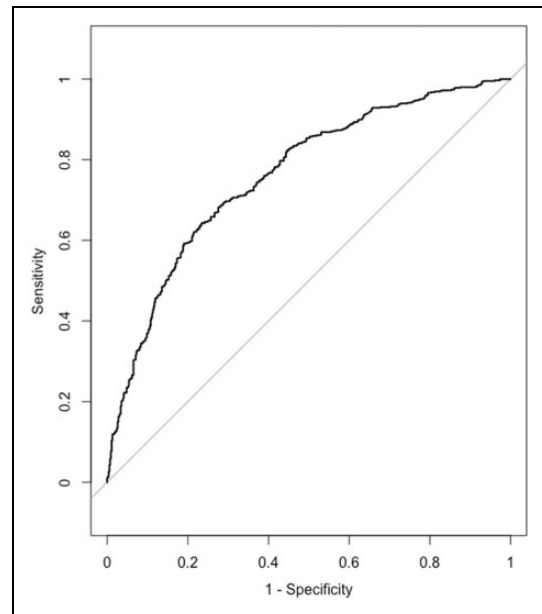
**Table 5.** Predictors and Odds Ratios (OR) in the Final Logistic Regression Model.

Predictor	OR	95% CI	Coding
Age	1.02	1.00-1.03	age in years
Duration of back pain before surgery	0.56	0.37-0.85	0 = less or equal to 1 year   1 = more than 1 year
Number of previous spine surgeries	0.65	0.44-0.97	0 = less or equal to 1   1 = more than 1
EQ5D-3L total score before surgery	0.12	0.08-0.16	0 = less or equal to 0.4   1 = more than 0.4
ODI total score before surgery	0.87	0.73-1.03	ODI total score
On sick leave before surgery	1.53	1.41-2.06	0 = yes   1 = no
On pain medication before surgery	0.69	0.45-1.08	0 = yes   1 = no
Fusion or disc prosthesis	2.4	1.71-3.38	0 = fusion   1 = disc prosthesis

Abbreviations: CI, confidence interval; ODI: Oswestry Disability Index.

<sup>a</sup>All increments are 1 except for the ODI total score before surgery—increment 15.

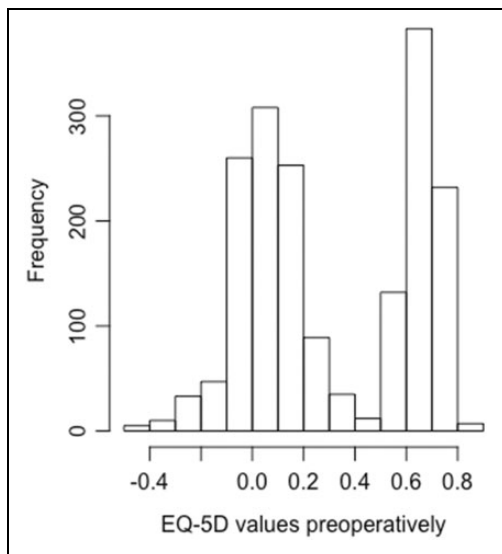
line with the previous Swedish study they reported decreased pain and disability in the surgical group compared with physiotherapy. In contrast, a small study from Norway of 64 patients comparing instrumented posterior fusion with rehabilitation detected no differences between groups at 12-month follow-up.<sup>22</sup> In the most recent randomized controlled trial by Fairbank et al.<sup>23</sup> A total of 349 patients were randomized to either surgery or a rehabilitation program. The mean ODI changed significantly in favor of surgery, but no significant differences between the treatment groups were observed in any of the other outcome measures. The main drawback in this study was a 28% crossover rate from the rehabilitation group



**Figure 2.** Receiver operating characteristic (ROC) curve for the final logistic regression model.

to the surgical group. Data analyses were carried out as an intention to treat analysis.

Loss to follow-up may bias the results. Two Scandinavian studies found that a loss to follow-up of as high as 22% did not bias conclusions of overall treatment effects.<sup>24,25</sup> The Norwegian study had a loss to follow-up of 28% and like the Danish follow-up study more males, younger patients and a higher percentage of smokers in the nonresponder group. In the Danish study, the nonresponders reported better outcomes than the responders.<sup>25</sup> Hence, it is reasonable to assume that this did not bias the results.



**Figure 3.** Histogram of EQ-5D values preoperatively.

The results of the logistic regression analysis for the fusion patients were somewhat surprising. The most important predictor was the EQ-5D-3L value before surgery. EQ-5D-3L total score before surgery was dichotomized with cut point 0.4. This value was chosen because this was the change point for the bimodal distribution of the EQ-5D-3L (Figure 3). The bimodal distribution of the EQ-5D-3L score is previously well described.<sup>26</sup> For a value of EQ-5D-3L score greater than 0.4 (preoperative value), there is approximately 8.6 times less chance of achieving a significant improvement in quality of life with a surgical intervention compared to those with a value less than 0.4. For each increase in ODI (patient deteriorated) by 15 points the chance of achieving a significant improvement in quality of life with surgical intervention increased approximately 1.2 times. The chance of improving quality of life after surgery decreased with increased duration of back pain, number of previous surgeries, and use of pain medication. The ROC curve showed acceptable discrimination with an AUC value of 0.75 (Figure 2). The cut point for the dependent variable (the dichotomized difference in the EQ-5D-3L score from before surgery to 1 year after surgery), which in this case equals the MCID, can be debated. We chose 0.19 (half the standard deviation). MCID as low as 0.14 or as high as 0.68 has been described for the EQ-5D-3L.<sup>27</sup> The minimal detectable change score for EQ-5D-3L at a 95% confidence interval were 0.43 in a study by Johnsen et al.<sup>28</sup> Using 0.43 as a cut point for the logistic regression analysis on our data will result in an unacceptable goodness-of-fit value for the Hosmer-Lemeshow test. Using the value of 0.43 or even more conservative measures would mean that the number of cases with a significant improvement of quality of life defined by this value would be less than or equal to 36% of all cases. We do not consider this realistic in a clinical setting, but regrettably, we do not have any anchor points in our data to back this up (example question: Were you satisfied with the results of surgery?). In lack of an

anchor point, it is recommended to use the value of half the standard deviation<sup>17,18</sup> as we did in this case. The MCID for the ODI is also a range rather than a well-defined cut point and in our opinion is not a better candidate for the dependent variable. However, it makes sense that with a relatively high EQ-5D-3L score before surgery or a correspondingly low ODI the chances of improving are less than in cases with ample opportunities for improvement. In other words, to be a candidate for spinal fusion the quality of life should be relatively low. However, it is a disadvantage to have undergone more than 1 previous surgery, to have back pain for more than 1 year before surgery, to take pain killers before the surgical intervention or to be on sick leave before surgery as this will decrease the chances of significantly improving the quality of life after spinal fusion. The use of disc prosthesis compared with spinal fusion improved the chances of a significant improvement in quality of life with a factor 2.4. However, the indications for disc prosthesis is not the same as for spinal fusion—and the patients selected for disc prosthesis differ in many ways from the patients selected for spinal fusion. Age as well as the preoperative total ODI score had only a small to moderate influence on the chance of achieving a significant improvement in quality of life after surgery—AUC decreased from 0.76 to 0.75 if age and ODI were excluded from the final model. It is noteworthy that smoking, body mass index, country, EQ5D-3L anxiety score before surgery, and number of levels were not included in the final model. Distress and depression have been known to influence the results after surgery, but the EQ5D-3L anxiety score is a rather weak predictor as shown by Carreon et al.<sup>29</sup> Only 57 of the 1184 patients were operated on more than 2 levels making it difficult to show a significant difference for this predictor. The direction of influence of the different predictors can be easily visualized using the dynamic nomogram (<https://dynamisknogramse.shinyapps.io/dynnomapp/>) and the characteristics of the patients can be entered giving the surgeons a hint of whether or not they should advocate surgery.

### Strength and Limitations

Register-based studies in general have not only advantages due to large sample sizes and high external validity but also limitations such as lower follow-up rates and inferior data quality compared with clinical trials. There is evidence in the literature that observational studies, correctly conducted according to the STROBE checklist, report results similar to randomized controlled trials.<sup>30</sup>

There are limitations with the current study design. The main limitations are the DDD diagnoses, which are assessed only by the operating surgeon. We have not been able to confirm the diagnoses in the registers but have relied on the treating surgeon registering the correct diagnosis in the register. Recent studies from both the Swespine and NORspine register showed that the diagnosis in the register and the surgical file was the same in 97% of the cases.<sup>31,32</sup>

Even though the Scandinavian countries are very comparable, cultural and language differences that could affect the

outcome of the same questionnaires translated in their own languages. Even if these have been cross-validated against other languages, we cannot exclude that this may have an impact on the results in this study.

## Conclusion

In this Nordic multicenter registry study, we found significant differences in preoperative patient characteristics, surgical treatment, and outcome. Danish patients were characterized by being older with higher preoperative NRS score than Swedish and Norwegian patients. Furthermore, the Danish patients had a higher rate of previous spine surgery and a lower preoperative EQ-5D-3L score. There were no differences in the ODI score between the countries. The most frequent surgical treatment was isolated fusion in all countries; however, the rate differed significantly with Sweden performing more TDR procedures than Denmark and Norway. Outcome was better in the TDR group, but the fusion group patients had higher frequency previous surgery, lower ODI score and higher EQ-5D-3L scores preoperatively with country-specific variations. Overall, there were statistically significant improvements in all outcomes measures with Norway having less improvement compared with Denmark and Sweden. Predictive modeling using logistic regression analysis resulted in a Shiny app decision support system (<https://dynamisknomogramse.shinyapps.io/dynnomapp/>) to help surgeons decide whether or not surgery should be advocated for a patient with a specific set of predictor variables.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## ORCID iD

Miranda van Hooff, MSc, PhD  <https://orcid.org/0000-0001-5313-6436>

## References

1. National Institute for Health and Care Excellence. Low back pain and sciatica in over 16s: assessment and management. NICE guideline [NG59]. <https://www.nice.org.uk/guidance/ng59>. Published November 2016. Accessed March 4, 2019.
2. Trotter JP. Patient registries: a new gold standard for “real world” research. *Ochsner J*. 2002;4:211-214.
3. Clement RC, Welander A, Stowell C, et al. A proposed set of metrics for standardized outcome reporting in the management of low back pain. *Acta Orthop*. 2015;86:523-533.
4. Nordic Medico-Statistical Committee 103. Health Statistics for the Nordic Countries. <http://norden.diva-portal.org/smash/get/diva2:874109/fulltext01.pdf>. Published November 2015. Accessed March 4, 2019.
5. Stromqvist B, Fritzell P, Hägg O, Jönsson B, Sandén B; Swedish Society of Spinal Surgeons. Swespine: the Swedish spine register: the 2012 report. *Eur Spine J*. 2013;22:953-974.
6. Solberg T, Olsen L. Yearly report from Norwegian Registry for Spine Surgery (NKR) 2014 (Årsrapport fra Nasjonalt kvalitetsregister for ryggkirurgi [NKR] 2014). [https://unn.no/Documents/Kvalitetsregistre/Nasjonalt%20kvalitetsregister%20for%20ryggkirurgi/%C3%85rsrapporter/%C3%A5rsrapport\\_NKR\\_2014.pdf](https://unn.no/Documents/Kvalitetsregistre/Nasjonalt%20kvalitetsregister%20for%20ryggkirurgi/%C3%85rsrapporter/%C3%A5rsrapport_NKR_2014.pdf). Published September 2015. Accessed December 9, 2017.
7. Sundhedsdata-Styrelsen. The Danish National Patient Registry. LPR advanced. <http://esundhed.dk/sundhedsregistre/LPR/Sider/LPR06A.aspx>. Accessed April 19, 2017.
8. Andersen M, Gehrchen M, Eiskjær OS; Danish Society of Spinal Surgery. DaneSpine Ryggkirurgi Årsrapport 2015. <http://drksdane.spine.dk/wm420129>. Published April 2016. Accessed April 19, 2017.
9. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine (Phila Pa 1976)*. 2000;25:2940-2952.
10. Hjerstad MJ, Fayers PM, Haugen DF, et al; European Palliative Care Research Collaborative (EPCRC). Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: a systematic literature review. *J Pain Symptom Manage*. 2011;41:1073-1093.
11. Dolan P. Modeling valuations for EuroQol health states. *Med Care*. 1997;35:1095-1108.
12. Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. *Pain*. 2011;152:2399-2404.
13. Briggs M, Closs JS. A descriptive study of the use of visual analogue scales and verbal rating scales for the assessment of postoperative pain in orthopedic patients. *J Pain Symptom Manage*. 1999;18:438-446.
14. Fairbank JC. Why are there different versions of the Oswestry Disability Index? *J Neurosurg Spine*. 2014;20:83-86.
15. Ostelo RW, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine (Phila Pa 1976)*. 2008;33:90-94.
16. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61:102-109.
17. Turner D, Schünemann HJ, Griffith LE, et al. The minimal detectable change cannot reliably replace the minimal important difference. *J Clin Epidemiol*. 2010;63:28-36.
18. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. 2nd ed. New York, NY: Wiley; 2000.
19. Berg S, Tropp H. Results from a randomized controlled study between total disc replacement and fusion compared with results from a spine register. *SAS J*. 2010;4:68-74. doi:10.1016/j.esas.2010.01.005
20. Möller H, Hedlund R. Surgery versus conservative treatment in adult isthmic spondylolisthesis—a prospective randomized study: part 1. *Spine (Phila Pa 1976)*. 2000;25:1171-1175.
21. Fritzell P, Hägg O, Wessburg P, Nordwall A; Swedish Lumbar Spine Study Group. Chronic back pain and fusion: a comparison of three surgical techniques: a prospective multicenter



- randomized study from the Swedish Lumbar Spine Study Group. *Spine (Phila Pa 1976)*. 2002;27:1131-1141.
22. Brox J, Sørensen R, Friis A, et al. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. *Spine (Phila Pa 1976)*. 2003;28:1913-1921.
  23. Fairbank J, Frost H, Wilson-MacDonald J, Yu LM, Barker K, Collins R; Spine Stabilisation Trial Group. Randomised controlled trial to compare surgical stabilisation of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: the MRC Spine Stabilisation trial. *BMJ*. 2005;330:1233. doi:10.1136/bmj.38441.620417.BF
  24. Solberg TK, Sørli A, Sjaavik K, Nygaard ØP, Ingebrigtsen T. Would loss to follow-up bias the outcome evaluation of patients operated for degenerative disorders of the lumbar spine? *Acta Orthop*. 2011;82:56-63.
  25. Højmark K, Støttrup C, Carreon L, Andersen MO. Patient-reported outcome measures unbiased by loss of follow-up. Single-center study based on DaneSpine, the Danish spine surgery registry. *Eur Spine J*. 2016;25:282-286.
  26. Parkin D, Devlin N, Feng Y. What determines the shape of an EQ-5D index distribution. *Med Decis Making*. 2016;36:941-951.
  27. Parker SL, Owoicho A, Paul AR, et al. Utility of minimum clinically important difference in assessing pain, disability and health status after transforaminal lumbar interbody fusion for degenerative lumbar spondylolisthesis. *J Neurosurg Spine*. 2011;14:598-604.
  28. Johnsen LG, Hellum C, Nygaard OP, et al. Comparison of the SF36D, the EQ5D, and the Oswestry Disability Index in patients with chronic low back pain and degenerative disc disease. *BMC Musculoskelet Disord*. 2013;14:218.
  29. Carreon LY, Djurasovic M, Dimar JR, et al. Can the anxiety domain of EQ-5D and mental health items from SF-36 help predict outcomes after surgery for lumbar degenerative disorders? *Neurosurg Spine*. 2016;25(3):352-356.
  30. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med*. 2000;342:1878-1886.
  31. Solberg T, Olsen L. Yearly report from Norwegian Registry for Spine Surgery (NKR) 2013 (Årsrapport fra Nasjonalt kvalitetsregister for ryggkirurgi [NKR] 2013). [https://unn.no/Documents/Kvalitetsregistre/Nasjonalt%20kvalitetsregister%20for%20ryggkirurgi/%C3%85rsrapporter/%C3%85rsrapport\\_NKR\\_2013\\_.pdf](https://unn.no/Documents/Kvalitetsregistre/Nasjonalt%20kvalitetsregister%20for%20ryggkirurgi/%C3%85rsrapporter/%C3%85rsrapport_NKR_2013_.pdf). Published October 2014. Accessed May 11, 2017.
  32. Endler P, Ekman P, Möller H, Gerdhem P. Outcomes of posterolateral fusion with and without instrumentation and of interbody fusion for isthmic spondylolisthesis: a prospective study. *J Bone Joint Surg Am*. 2017;99:743-752.