Lumbar spinal stenosis

comparison of surgical practice variation and clinical outcome in three national spine registries

Lønne, Greger; Fritzell, Peter; Hägg, Olle; Nordvall, Dennis; Gerdhem, Paul; Lagerbäck, Tobias; Andersen, Mikkel; Eiskjaer, Søren; Gehrchen, Martin; Jacobs, Wilco; van Hooff, Miranda L; Solberg, Tore K

Published in:
Spine Journal

DOI (link to publication from Publisher):
10.1016/j.spinee.2018.05.028

Publication date:
2019

Document Version
Accepted author manuscript, peer reviewed version

Citation for published version (APA):
Lønne, G., Fritzell, P., Hägg, O., Nordvall, D., Gerdhem, P., Lagerbäck, T., Andersen, M., Eiskjaer, S., Gehrchen, M., Jacobs, W., van Hooff, M. L., & Solberg, T. K. (2019). Lumbar spinal stenosis: comparison of surgical practice variation and clinical outcome in three national spine registries. *Spine Journal*, *19*(1), 41-49. https://doi.org/10.1016/j.spinee.2018.05.028

Accepted Manuscript

Title: Lumbar spinal stenosis: comparison of surgical practice variation and clinical outcome in three national spine registries

Author: Greger Lønne, Peter Fritzell, Olle Hägg, Dennis Nordvall, Paul Gerdhem, Tobias Lagerbäck, Mikkel Andersen, Søren Eiskjaer, Martin Gehrchen, Wilco Jacobs, Miranda L. van Hooff, Tore K. Solberg

PII: S1529-9430(18)30253-5
DOI: https://doi.org/10.1016/j.spinee.2018.05.028
Reference: SPINEE 57694

To appear in: *The Spine Journal*

Received date: 16-1-2018
Revised date: 16-5-2018
Accepted date: 17-5-2018

Lumbar spinal stenosis: Comparison of surgical practice variation and clinical outcome in three national spine registries

Greger Lønne[a,b,c], M.D., Ph.D., Peter Fritzell[d,e,f,g], M.D., Ph.D., Olle Hägg[h,i], M.D., Ph.D., Dennis Nordvall[g], Paul Gerdhem[j,k], M.D., Ph.D., Tobias Lagerbäck[j,k], M.D., Mikkel Andersen[l], M.D., Søren Eiskjaer[m], M.D., Martin Gehrchen[n], M.D., Ph.D., Wilco Jacobs[o], M.Sc., Ph.D., Miranda L van Hooff[p,q], M.Sc., Ph.D., Tore K Solberg[c,r,s], M.D., Ph.D.,

[a]Department of Orthopaedics, Innlandet Hospital Trust, Lillehammer, Norway
[b]National Advisory Unit on Spinal Surgery, St. Olavs Hospital, Trondheim University Hospital, Trondheim, Norway
[c]The Norwegian Registry for Spine Surgery (NORspine), Northern Norway Regional Health Authority, Norway
[d]Department of Orthopaedics, Capio St Göran Hospital, Stockholm, Sweden
[e]Department of Surgical Sciences, Division of Orthopaedics, Uppsala University, Sweden
[f]Strömstad akademi, Strömstad, Sweden
[g]Qulturum Center for Learning and Innovation in Healthcare, Jönköping, Sweden
[h]Spine Center Göteborg, Göteborg, Sweden
[i]Swespine Steering Group, Swedish National Spine Register, Jönköping, Sweden
[j]Department of Orthopaedics, Karolinska University Hospital Huddinge, Sweden
[k]Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Stockholm, Sweden
[l]Sector for Spine Surgery and Research, Lillebaelt Hospital, Middelfart, Denmark
[m]Department of Orthopedic Surgery, Aalborg University Hospital, Aalborg, Denmark
[n]Spine Unit, Department of Orthopaedic Surgery, Rigshospitalet University of Copenhagen, Denmark
[o]The Health Scientist, The Hague, The Netherlands
[p]Department Research, Sint Maartenskliniek, Nijmegen, The Netherlands
[q]Department of Orthopedics, Radboud university medical center, Nijmegen, The Netherlands
[r]Department of Neurosurgery, University Hospital of Northern Norway, Tromsø, Norway
[s]Institute of Clinical Medicine, University of Tromsø The Arctic University of Norway, Tromsø, Norway

Corresponding author:
Greger Lønne
Innlandet Hospital Trust
Department of orthopedic surgery
Anders Sandvigs gt. 17
2629 Lillehammer
Norway

Mob: +47 971 14 107
Email: gloenne@mac.com

ABSTRACT

Background: Decompression surgery for lumbar spinal stenosis (LSS) is the most common spinal procedure in the elderly. To avoid persisting low back pain, adding arthrodesis has been recommended, especially if there is a coexisting degenerative spondylolisthesis. However, this strategy remains controversial, resulting in practice-based variation.

Purpose: To evaluate in a pragmatic study if surgical selection criteria and variation in use of arthrodesis in three Scandinavian countries can be linked to variation in treatment effectiveness.

Study design: An observational study based on a combined cohort from the national spine registries of Norway, Sweden, and Denmark.

Patient Sample: Patients aged 50 and higher operated 2011–2013 for LSS were included.

Outcome Measures: Patient-reported outcome measures (PROMs) Oswestry disability index (ODI) (primary outcome), numeric rating scale (NRS) for leg pain and back pain, and health-related quality of life (EQ-5D). Analysis included case-mix adjustment. In addition, we report differences in hospital stay.

Methods: Analyses of baseline data were done by analysis of variance (ANOVA), Chi-square, or logistic regression tests. The comparisons of the mean changes of PROMs at one-year follow-up between the countries were done by ANOVA (crude) and analyses of covariance (ANCOVA, case mix adjustment). There are no conflicts of interest. Funding was received from the Danish Society of Spinal Surgery ($5,925), the Northern Norway Regional Health Authority ($5,925) and from Swedish Association of Local Authorities and Regions ($11,885). The sponsor had no role in the acquisition of data, analysis, or preparation of the manuscript.

Results: Out of 14,223 included patients, 10,890 (77%) responded at one-year follow-up. Apart from fewer smokers in Sweden and higher comorbidity rate in Norway, baseline characteristics were similar. The rate of additional fusion surgery (patients without, with spondylolisthesis) was: Norway 11% (4%, 47%), Sweden 21% (9%, 56%) and Denmark 28% (15%, 88%). At one-year follow-up the mean improvement for ODI (95%CI) was: Norway 18

(17 to 18), Sweden 17 (17 to 18), and Denmark 18 (17 to 19). Patients operated with arthrodesis had prolonged hospital stay.

Conclusions: Real life data from three national spine registers showed similar indications for decompression surgery, but significant differences in the use of concomitant arthrodesis in Scandinavia. Additional arthrodesis was not associated with better treatment effectiveness.

Keywords: lumbar spinal stenosis, spine registry, decompressive surgery, case mix adjustment, spine fusion, spine arthrodesis

INTRODUCTION

Low back pain is the leading specific cause for years lived with disability worldwide [1]. Narrowing of the spinal canal, known as lumbar spinal stenosis (LSS) is the most common indication for spine surgery in the elderly population. LSS typically causes symptoms of low back pain, lower extremity pain and numbness due to nerve root compression, resulting in walking disability [2]. Decompression of the spinal canal is the key objective of surgery and is considered superior to non-surgical treatment for patients with moderate to severe LSS [3]. Often, there is a coexisting degenerative spondylolisthesis, i.e. a slip of one vertebra in relation to another. Traditionally, this radiological finding has been regarded as a sign of segmental instability. Although this interpretation has been disputed, adding surgical fusion between the two vertebrae (arthrodesis) in addition to decompression has been recommended to prevent persisting back pain [4, 5]. However, several recent studies found no effect of additional arthrodesis surgery [6-8]. Due to lack of uniform guidelines in this field, there is a large and possibly unwarranted practice variation in the use of additional arthrodesis [9, 10]. In a recent study fusion rate (with, without spondylolisthesis) was considerably lower in university hospitals of Norway (44%, 6%) compare to Boston, US (95%, 29%) [11]. In the

US, rising costs connected to arthrodesis of the lumbar spine have attracted the attention of health providers and policy makers. In 2011 spinal fusion accounted for the highest aggregate hospital costs of any surgical procedure performed in U.S. hospitals ($12.8 billion) [12]. The higher cost connected to arthrodesis surgery should be justified by better patient-reported outcome. In 2015, the International consortium for health outcome measurement (ICHOM) recommended a set of patient-reported outcome measures (PROMs) for evaluating surgical treatment of degenerative conditions in the lumbar spine to facilitate clinical studies across nations and centers [13]. The national spine surgery registries of Norway, Sweden, and Denmark were among the collaborators. Scandinavian countries are characterized by a genetically homogenous population, similar social security systems, and public based health care and health insurance systems, facilitating comparative studies [14]. The incidence of surgically treated lumbar spinal stenosis is similar (30-35/100 000/year) based on imputed numbers from the registries. Clinical registries collecting data from everyday practice can evaluate different treatment strategies by linking practice-based variation to patient-reported outcomes in a pragmatic trail. Unlike randomized controlled trials, registry-based studies allow for surgeons and patients preferences to be included in the process prior to surgery, as in the "real world" of clinical practice, and adds external validity to already published data from randomized controlled trails [15]. Such information may aid in guideline development and resource allocation.

The aims of this observational multinational register study were to compare practice-based variation in surgical treatment of LSS by; (1) surgical selection criteria (preoperative patient characteristics), (2) type of surgery (decompression only or decompression plus arthrodesis), and (3) to assess if practice-based variations were associated to different patient-reported outcomes (crude and case mix adjusted), in a large combined registry cohort from three Scandinavian countries.

METHODS

This observational study reviews data from the national spine registries of Norway (NORspine), Sweden (Swespine), and Denmark (DaneSpine). Eligible patients were aged 50 or older with no history of previous lumbar spine surgery, operated for LSS during 2011, 2012, or 2013. At baseline, the surgeon recorded diagnosis and treatment according to standardized questionnaire. The diagnosis of LSS was based on the surgeons' clinical judgment and assessment of magnetic resonance imaging, MRI. Concomitant spondylolisthesis is defined as a visible slip, 3 mm or more, of one vertebra in relation to another. All patients received surgical decompression, some with concomitant arthrodesis.

The registers

All three national spine registries are designed for quality control and research. The participation is voluntary for the surgical departments as well as the patient. At admission for surgery (baseline), the patient reports data on demographics, risk factors, and PROMs. During the hospital stay, the surgeon records diagnosis, type of surgery, and perioperative complications. At one-year follow-up, questionnaires are distributed from the central national registry office, completed at home by the patients, and returned in pre-stamped envelopes. The treating hospitals are not involved in follow up. The oldest registry, Swespine, has included patients since 1998. Swespine covers approximately 95% of the surgical units in Sweden. Completeness, the proportion of operated patients reported to Swespine, was approximately 75%. NORspine is based on the concept of Swespine, and was founded in 2007 (coverage 95%, completeness of 65%). DaneSpine was acquired by the Danish Spine Society from the Swedish Society of Spinal Surgeons in 2009 and has successively been implemented (coverage 80%, completeness 62%).

Patient-reported outcome measures (PROMs)

We used the ICHOM recommended set of PROMs [13]. The primary outcome was the Oswestry Disability Index (ODI, version 2.1), a standard for measuring back pain related disability, ranging from 0 (no disability) to 100 (bedridden) [16].

Secondary outcome measures were numeric rating scales (NRS) for back and leg pain, ranging from 0 (no pain) to 10 (worst conceivable pain). Health-related quality of life was measured with the Euro-Qol-5D (EQ-5D) ranging from -0.596 to 1, with higher scores indicating better quality of life.

NORspine used the NRS for leg and back pain, while Swespine and DaneSpine used the Visual Analogue Scale (VAS), ranging from 0-100. Conversion to NRS was done by dividing the VAS score by ten with a stochastic approximation of decimals to the closest integer.

Data handling and analysis

Anonymous data from the three registers were pooled and stored on the Swespine data server. Missing or out of range data on gender, age, height, or weight were deleted (Figure 1). In case of missing outcome data case exclusion analysis by analysis was used. Furthermore, cases with missing date of surgery and follow-up were excluded.

Comparisons of data

Analysis of baseline data included PROM-scores, age at date of surgery, gender, height, weight, smoking habits, sick leave, and duration of leg and back pain presented as mean (95% confidence interval), or number (%). Variables were analyzed by analysis of variance (ANOVA), Chi-square, or logistic regression tests. The comparisons of the mean changes of PROMs at one-year follow-up were done by ANOVA (crude) and analyses of covariance (ANCOVA, case mix adjustment). The minimal clinically important change (MCIC) is the minimal PROM score change that is perceived as meaningful by individual patients, irrespective of statistical significance level. The MCIC was defined as 15 for ODI and 2.0 for NRS back pain and leg pain, and 0.15 for EQ-5D [17-19] within groups. To compare clinically meaningful differences in outcomes between groups, we compared the percentage of patients achieving at least 30% improvement of ODI, and NRS back pain and leg pain [18]. The absolute 12 months follow up score defining a patient acceptable symptom state was set to ODI $\leq$ 22 [20].

Sample size

Due to the large sample size (n>10,000), ODI differences as small as 2 points between the groups would be reported significant (power 90%, significance level 5%), i.e. far below what is considered as clinically relevant [16].

Non-response analysis

A non-response analysis was performed by comparing all available baseline variables between those who responded to the one-year follow-up with those who did not.

Ethics

This study was approved by ethical review boards in Norway (REC South-east B: 2014/2219), Sweden (Dnr 2015/181-31), and Denmark (Projekt-ID: S-20160091). It was conducted and reported in accordance with the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) checklist and the study protocol, available at clinicaltrails.gov (ID: NCT02897947).

Funding

Funding was received from the Danish Society of Spinal Surgery ($5,925), the Northern Norway Regional Health Authority ($5,925) and from Swedish Association of Local

Authorities and Regions ($11,885). The funding sources had no role in the study design, analysis, and interpretation of data, in the writing of the report, and in the decision to submit the paper for publication.

Conflict of interest

All authors declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

RESULTS

Baseline

At baseline 14223 were included (Norway: n =3173, Sweden: n = 7389, and Denmark: n = 3661). At one-year follow-up, 10890 (77%) responded (Norway: n =2559 (81%), Sweden: n = 5990 (81%), and Denmark: n = 2341 (64%)). Figure 1 shows the exclusion flowchart. Gender, age, and BMI were similar in the three countries. Fewer were smoking in Sweden, and a higher comorbidity rate was found in Norway (Table 1).

Mean baseline disability (ODI (95%CI)) was slightly worse in Sweden (44 (43 to 44)), compared to Denmark (41 (40 to 41)) and Norway (40 (39 to 40)). Health related quality of life (EQ-5D ±SD) was better in Denmark (0.42 (0.41 to 0.43) vs. Sweden (0.37 (0.36 to 0.38)) and Norway (0.37 (0.36 to 0.39)). Accordingly, NRS leg pain and back pain intensity were less in Denmark. In this study, the non-responders at one year follow-up (n= 3333) were slightly younger and more often smokers, but otherwise similar to the responders (n=10890) at baseline (Table 2). Multiple levels surgery (two, three levels) was less frequent in Norway (30%, 6%), than in Sweden (34%, 12%) and Denmark (35%, 12%).

Rate of concomitant arthrodesis

The rate of concomitant arthrodesis was significantly different between the three countries: Norway 11%, Sweden 21%, and Denmark 28%. For the subgroup of patients with concomitant spondylolisthesis, the rate of arthrodesis was higher: Norway 47%, Sweden 56% and Denmark 88% (Figure 2).

Perioperative complications and differences in days at hospital

The frequency of dural tear was: Norway 4.8%, Sweden 5.7%, and Denmark 5.3%, p=0.088. The frequency of excessive bleeding was: Norway 0.16%, Sweden 0.45%, and Denmark 0.30%, p=0.058 the frequency of nerve root injury was: Norway 0.16%, Sweden 0.09%, and Denmark 0.03%, p=0.204. The overall rate of perioperative complications was: Norway 5.5%, Sweden 6.2%, and Denmark 5.0%, p=0.033. In the combined cohorts the in hospital surgeon reported complication rate was 5.8% for both decompression only and decompression with additional arthrodesis.

In Norway, the mean number of days at hospital (SD) (day 1; day of admission) for patients operated with decompression alone compared to decompression plus arthrodesis was 3.0 (2.8) vs. 7.3 (3.9). In Sweden (day 1; day of admission) the corresponding numbers were 3.6 (3.5) vs. 5.3 (3.1) and in Denmark (day 1; day of operation) 2.0 (1.8) vs. 4.7 (3.2).

Outcome at one year

At one-year follow-up the mean improvement for ODI (95%CI) was in Norway 18 (17 to 18), Sweden 17 (17 to 18) and Denmark 18 (17 to 19). In the case-mix analysis adjusted for age, gender, BMI, smoking, any comorbidity, and the preoperative value of ODI, the corresponding values were 16 (16 to 17), 18 (18 to 19), and 17 (17 to 17) (Table 3). A MCIC-value of 30% improvement of ODI was achieved by 58% in Norway, 53% in Sweden and 50% in Denmark, ($p<0.001$). ODI score 22 or below was achieved by 64% in Norway, 64% in Sweden and 64% in Denmark ($p=0.837$). There were no differences in rate of patients reaching MCIC for leg pain or back pain between the countries (Table 3).

Subgroup analysis

In the combined cohort, patients operated for LSS without spondylolisthesis had an unadjusted mean ODI improvement (95%CI) of 17 (17 to 18) in the decompression only group and 18 (17 to 20) in the decompression plus arthrodesis group. Using case mix adjusted analyses the corresponding numbers were 17 (17 to 18) and 19 (18 to 20). For patients with a concomitant spondylolisthesis, the improvement in unadjusted mean (95%CI) was 17 (16 to 18) in the decompression only group and 20 (19 to 21) in the decompression and arthrodesis group. Corresponding case mixed values were 17 (17 to 18) and 18 (18 to 19) (Table 4). When comparing outcomes of patients with and without spondylolisthesis between the three countries, no clinically relevant differences were found (Figure 3).

DISCUSSION

To our knowledge, this represents the worlds' largest observational study of patients operated for LSS, and the first comparison across countries using the ICHOM-recommended core data set. Even though the selection criteria for surgery in terms of demographic characteristics, pain intensity and disability were similar, we found a significant practice variation, i.e. use of additional arthrodesis surgery was almost three times higher in Denmark and two times higher in Sweden as compared to Norway (Figure 2). This demonstrates that even in homogenous populations with similar health care systems the treatment traditions can vary considerably.

We observed longer hospital stay among patients operated with additional arthrodesis, which, together with the implants used, indicates higher cost but no better treatment effectiveness.

Our findings are in accordance with a recent Swedish randomized controlled trial (RCT) by Försth et al. of 247 patients showing that additional arthrodesis neither reduced reoperation rates nor improved clinical outcomes (ODI) [6]. A randomized controlled trial from the US by Ghogawala et al. involving 66 patients found that additional arthrodesis surgery for LSS with mild spondylolisthesis reduced the risk for reoperation and gave larger improvement of physical health–related quality of life (generic SF 36) than laminectomy alone [7]. For all other outcomes, including the disease specific ODI, no difference was found. This study has been heavily criticized, also because reoperation rate during follow-up was remarkably high [21]. Higher frequency of reoperations in the US may however reflect potential cultural differences in patient expectations, difference in treatment traditions and incentives for arthrodesis surgery driven by health insurance and reimbursement programs compare to those found in countries like Sweden.

A Swedish non-randomized registry study of 5390 LSS patients with or without spondylolisthesis operated between 1998 and 2008, found no benefit of additional arthrodesis after two years [8]. Similar results were shown in a Swiss multicenter study from 2017 of 185 patients with LSS and spondylolisthesis after three years [22]. A recent Norwegian pragmatic comparative effectiveness study showed marginally better improvement (less than MCIC), of back pain among LSS patients with spondylolisthesis receiving decompression plus arthrodesis. No such association was found for ODI [23].

We also found a large difference in the use of additional arthrodesis in patients without spondylolisthesis in 2011 – 2013. This treatment strategy has been discussed among spinal surgeons for many years, and is not in accordance with guidelines from 2013, where "decompression alone is suggested for patients with leg predominant symptoms without instability" [2, 4, 9]. The term "instability" is poorly defined, but has been linked to low back pain, a frequent symptom in LSS. This may explain the practice variation, also shown in a previous study where the arthrodesis rate in cases without spondylolisthesis was 29% in Boston (US), compared to only 6% in Norway [11]. We observed a rising rate of arthrodesis from Norway, via Sweden, to Denmark across the countries (Figure 2), but no corresponding trend (dose-response effect) in terms of higher treatment effectiveness (Table 3). In fact, the mean improvement of back pain in the spondylolisthesis group was somewhat higher in Norway (3.6) than in Denmark (2.7), which had the highest rate of arthrodesis (Figure 3). Hence, this study does not support the argument that arthrodesis prevents low back pain related to instability in spinal stenosis patients. The different frequency of multiple level surgery was small, and can neither explain the difference in the fusion rate, nor the lack of difference in outcome.

We did both crude analysis and case mix analysis. Crude data shows small, not clinical relevant difference in the outcome between those with spondylolisthesis having decompression and fusion, but these differences vanished after the case mix adjustment (Table 4).

Fox et al. concluded in 1996 that radiological instability was common after decompression for degenerative LSS without spondylolisthesis, but correlated poorly with clinical outcome (back pain) [24]. The quality of some earlier studies advocating additional arthrodesis routinely is low due to small sample sizes, weak design, and outcome based on radiological findings [25]. Moreover, a change towards using more minimally invasive decompression techniques may have reduced the risk for postoperative instability [26]. Previous studies show that arthrodesis adds higher risk of major complications, and even mortality [27]. Like Ghogawala et al., we found no association between the use of concomitant arthrodesis and surgeon reported complications [7].

Comorbidity rate in NORspine was physician-reported and higher compared to the patient-reported rate in Swespine and DaneSpine. However, outcomes were similar, also when adjusting for comorbidity (Table 3). Between countries with larger diversity in demographic, socio-economic and cultural features, case mix adjustment may be more important.

Even if the differences in effects sizes were smaller than considered as clinically relevant, subgroups of patients may benefit from additional arthrodesis. This should be investigated further in studies utilizing more precise data on radiological findings and with long term follow-up to assess reoperation rates.

Quality assurance

Loss to follow-up may bias the results. Two Scandinavian studies found that a loss to follow-up of as high as 23% would not bias conclusions about overall treatment effects [28, 29]. They found, similar to our results, that non-responders were younger and more likely smokers. Therefore, it would be reasonable to assume that loss to follow up did not bias our results.

Strength and limitations

Register-based studies in general have advantages such as large sample sizes and high external validity, but also limitations due to lack of randomization, lower follow-up rates, and lower internal validity compared to closely monitored clinical trials. In contrast to RCTs, this study allows surgeons and patients preferences to be included in a shared decision-making process prior to surgery, like in the "real world" of clinical practice. Still, there is increasing evidence in the literature that observational studies, conducted according to STROBE check list, report corresponding results similar to those found in RCTs [30].

There are limitations associated with this work. Even though registry data were collected prospectively for quality control and research, the hypotheses were decided on in retrospect. In addition, we did not have exact data on reoperation rates and only one-year follow-up. Reoperation rates may be as high as 20% at long term (3 to 5 years) [6], but previous studies have shown that clinical outcomes are stable up to 5 years [6].

"In Scandinavia it is recommended to try conservative treatment prior to surgery for lumbar spinal stenosis. Previous studies show that the content of non-operative care is hard to define [31], and the effects of different conservative treatment alternatives are ambiguous. Since no uniform Scandinavian guidelines for such treatment exist, the type of preoperative conservative treatment was not recorded in the registries, only duration of symptoms.

The use of the ICHOM concept and adding case mix analyses makes comparisons more credible, but a relative small set of baseline variables has been used for case mix adjustment.

CONCLUSION

Real life data from three national spine registers showed similar indications for decompression surgery, but significant differences in the use of concomitant arthrodesis in Scandinavia. Additional arthrodesis was not associated with better treatment effectiveness.

ACKNOWLEDGEMENT

The authors thank all the patients and surgeons contributing with data to the spine registers in Sweden, Denmark and Norway.

References

1. Vos T, Flaxman AD, Naghavi M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2163-96.
2. Kreiner DS, Shaffer WO, Baisden JL, et al. An evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spinal stenosis (update). Spine J. 2013;13(7):734-43.
3. Jacobs W, Rubinstein S, Willems P, et al. The evidence on surgical interventions for low backdisorders, an overview of systematic reviews. . Eur Spine J. 2013;22:1936–39.
4. Peul WC, Moojen WA. Fusion for Lumbar Spinal Stenosis--Safeguard or Superfluous Surgical Implant? The New England journal of medicine. 2016;374(15):1478-9.
5. Schroeder GD, Kepler CK, Kurd MF, et al. Rationale for the Surgical Treatment of Lumbar Degenerative Spondylolisthesis. Spine (Phila Pa 1976). 2015;40(21):E1161-6.
6. Forsth P, Olafsson G, Carlsson T, et al. A Randomized, Controlled Trial of Fusion Surgery for Lumbar Spinal Stenosis. The New England journal of medicine. 2016;374(15):1413-23.
7. Ghogawala Z, Dziura J, Butler WE, et al. Laminectomy plus Fusion versus Laminectomy Alone for Lumbar Spondylolisthesis. The New England journal of medicine. 2016;374(15):1424-34.
8. Forsth P, Michaelsson K, Sanden B. Does fusion improve the outcome after decompressive surgery for lumbar spinal stenosis?: A two-year follow-up study involving 5390 patients. The bone & joint journal. 2013;95-B(7):960-5.
9. Deyo RA. Commentary: Clinical practice guidelines: trust them or trash them? Spine J. 2013;13(7):744-6.
10. Kepler CK, Vaccaro AR, Hilibrand AS, et al. National trends in the use of fusion techniques to treat degenerative spondylolisthesis. Spine (Phila Pa 1976). 2014;39(19):1584-9.
11. Lønne G, Schoenfeld AJ, Cha TD, Nygaard ØP, Zwart JA, Solberg TK. Variation in selection criteria and approaches to surgery for Lumbar Spinal Stenosis among patients treated in Boston and Norway. Clinical Neurology and Neurosurgery. 2017;156:77-82.
12. Weiss A, Elixhauser A, Andrews R. Characteristics of Operating Room Procedures in U.S. Hospitals, 2011. HCUP Statistical Brief #170 Agency for Healthcare Research and Quality, Rockville, MD. 2014.
13. 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. 2015:1-11.
14. NOMESCO. Health Statistics for the Nordic Countries. 2015;103.
15. Patsopoulos NA. A pragmatic view on pragmatic trials. Dialogues Clin Neurosci. 2011;13(2):217-24.
16. Fairbank JC, Pynsent PB. The Oswestry Disability Index. Spine. 2000;25(22):2940-52-discussion 52.
17. Fairbank JC. Why are there different versions of the Oswestry Disability Index? Journal of neurosurgery Spine. 2014;20(1):83-6.

18. 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(1):90-4.

19. Parker SL, Godil SS, Shau DN, Mendenhall SK, McGirt MJ. Assessment of the minimum clinically important difference in pain, disability, and quality of life after anterior cervical discectomy and fusion: clinical article. Journal of neurosurgery Spine. 2013;18(2):154-60.

20. van Hooff ML, Mannion AF, Staub LP, Ostelo RW, Fairbank JC. 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. Spine J. 2016;16(10):1221-30.

21. Epstein NE. Commentary on: Laminectomy plus fusion versus laminectomy alone for lumbar spondylolisthesis by Ghogawala Z, Dziura J, Butler WE, Dai F, Terrin N, Magge SN, et al. NEJM 2016;374 (15):1424-34. Surg Neurol Int. 2016;7(Suppl 25):S644-S7.

22. Ulrich NH, Burgstaller JM, Pichierri G, et al. Decompression Surgery Alone Versus Decompression Plus Fusion in Symptomatic Lumbar Spinal Stenosis: A Swiss Prospective Multi-center Cohort Study with 3 Years of Follow-up. Spine (Phila Pa 1976). 2017.

23. Austevoll IM, Gjestad R, Brox JI, et al. The effectiveness of decompression alone compared with additional fusion for lumbar spinal stenosis with degenerative spondylolisthesis: a pragmatic comparative non-inferiority observational study from the Norwegian Registry for Spine Surgery. Eur Spine J. 2016.

24. Fox MW, Onofrio BM, Onofrio BM, Hanssen AD. Clinical outcomes and radiological instability following decompressive lumbar laminectomy for degenerative spinal stenosis: a comparison of patients undergoing concomitant arthrodesis versus decompression alone. J Neurosurg. 1996;85(5):793-802.

25. Herkowitz HN, Kurz LT. Degenerative lumbar spondylolisthesis with spinal stenosis. A prospective study comparing decompression with decompression and intertransverse process arthrodesis. J Bone Joint Surg Am. 1991;73(6):802-8.

26. Nerland US, Jakola AS, Solheim O, et al. Minimally invasive decompression versus open laminectomy for central stenosis of the lumbar spine: pragmatic comparative effectiveness study. BMJ. 2015;350:h1603.

27. Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG. Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. JAMA. 2010;303(13):1259-65.

28. Solberg TK, Sorlie A, Sjaavik K, Nygaard OP, 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(1):56-63.

29. Hojmark K, Stottrup 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(1):282-6.

30. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. New England Journal of Medicine. 2000;342(25):1878-86.

31. Weinstein JN, Tosteson TD, Lurie JD, Tosteson A. Surgical versus Non-Operative Treatment for Lumbar Spinal Stenosis Four-Year Results of the Spine Patient Outcomes Research Trial (SPORT). Spine. 2010:1.

LEGENDS

Figure 1: Flowchart showing study enrolment.

Figure 2: Four bar charts showing rate of arthrodesis in lumbar spinal stenosis with or without spondylolisthesis in Norway, Sweden, and Denmark.

Figure 3: Four bar charts showing the mean improvement in Oswestry Disability Index (ODI), Numeric rate score (NRS) leg pain, NRS back pain, and Euro-Qual – Five Dimensions (EQ-5D) in patients without and with spondylolisthesis in Norway, Sweden, and Denmark.

1 **Table 1 Baseline characteristics of patients operated in Norway, Sweden, and Denmark**

Total (n = 14223)	Norway (n =3173)	Sweden (n = 7389)	Denmark (n = 3661)	(P-value)
Age, years (SD)	67.5 (9.0)	68.9 (8.9)	68.6 (9.1)	<0.001*
BMI, kg/m^2 (SD)	27.3 (4.3)	27.4 (4.1)	27.1 (4.4)	0.002*
Females, n (%)	1701 (54%)	4075 (54%)	2006 (55%)	0.595†
Smokers, n (%)	660 (21.1%)	678 (9.3%)	792 (22.0%)	<0.001†
Any comorbidity, n (%)	804 (25.4%)	591 (8.0%)	352 (9.6%)	<0.001†
Neurological comorbidity, n (%)	76 (2.4%)	201 (2.8%)	84 (2.3%)	0.334†
Heart comorbidity, n (%)	686 (21.7%)	313 (4.3%)	201 (5.4%)	<0.001†
Cancer comorbidity, n (%)	91 (2.9%)	77 (1.0%)	67 (1.8%)	<0.001†
Preoperative PROM				
ODI (95%CI)	40 (39 to 40)	44 (43 to 44)	41 (40 to 41)	<0.001*
NRS leg pain (95%CI)	6.6 (6.5 to 6.7)	6.4 (6.3 to 6.4)	5.8 (5.7 to 5.9)	<0.001*
NRS back pain (95%CI)	6.4 (6.4 to 6.5)	5.6 (5.6 to 5.7)	5.1 (4.9 to 5.2)	<0.001*
EQ-5D (95%CI)	0.37 (0.36 to 0.39)	0.37 (0.36 to 0.38)	0.42 (0.41 to 0.43)	<0.001*
Preoperative duration of leg pain >1 year, n (%)	3173 (68%)	4996 (68%)	3661 (55%)	<0.001†
Preoperative duration of back pain >1 year, n (%)	2470 (78%)	5567 (75%)	2473 (68%)	<0.001†

2 *ODI = Oswestry Disability Index, NRS = Numeric Rating Scale, EQ-5D = Euro-Qol-5D 3 levels, BMI = Body Mass Index, 95%CI = 95% confidence interval*

3 * *ANOVA F-test*

4

5 † *Pearson's Chi-square test*

6

7

8

9

Table 2 - Baseline characteristics of responders and non-responders

Total (n = 14223)	Responders n = 10890 (77%)	Non-responders n = 3333 (23%)	P value
Age, years (SD)	68.7 (8.8)	67.9 (9.7)	<0.001*
BMI, kg/m2 (SD)	27.3 (4.2)	27.4 (4.6)	0.297*
Females, n (%)	5905 (54%)	1827 (55%)	0.360†
Smokers, n (%)	1449 (14%)	681 (21%)	<0.001†
Any co-morbidity, n (%)	1299 (12%)	448 (13%)	0.020†
Neurological co-morbidity, n (%)	265 (2.4%)	96 (2.9%)	0.151†
Heart co-morbidity, n (%)	910 (8.4%)	290 (8.7%)	0.531†
Cancer co-morbidity, n (%)	166 (1.5%)	69 (2.1%)	0.031†
Preoperative duration of leg pain >1 year, n (%)	7017 (64%)	2164 (65%)	0.604†
Preoperative duration of back pain >1 year, n (%)	8032 (74%)	2478 (74%)	0.496†
Preoperative PROM values			
ODI (SD)	40 (16)	43 (16)	<0.001*
NRS leg pain (SD)	6.3 (2.5)	6.2 (2.6)	0.112*
NRS back pain (SD)	5.6 (2.7)	5.7 (2.7)	0.205*
EQ-5D (SD)	0.40 (0.31)	0.36 (0.32)	<0.001*

Data are shown as mean (SD), or number (%). P-values for comparison between responders and non-responders are shown. PROM = Patients Reported Outcome Measures.

Student's t-test

†*Pearson's Chi-square test*

Table 3 Improvements in PROMs by country

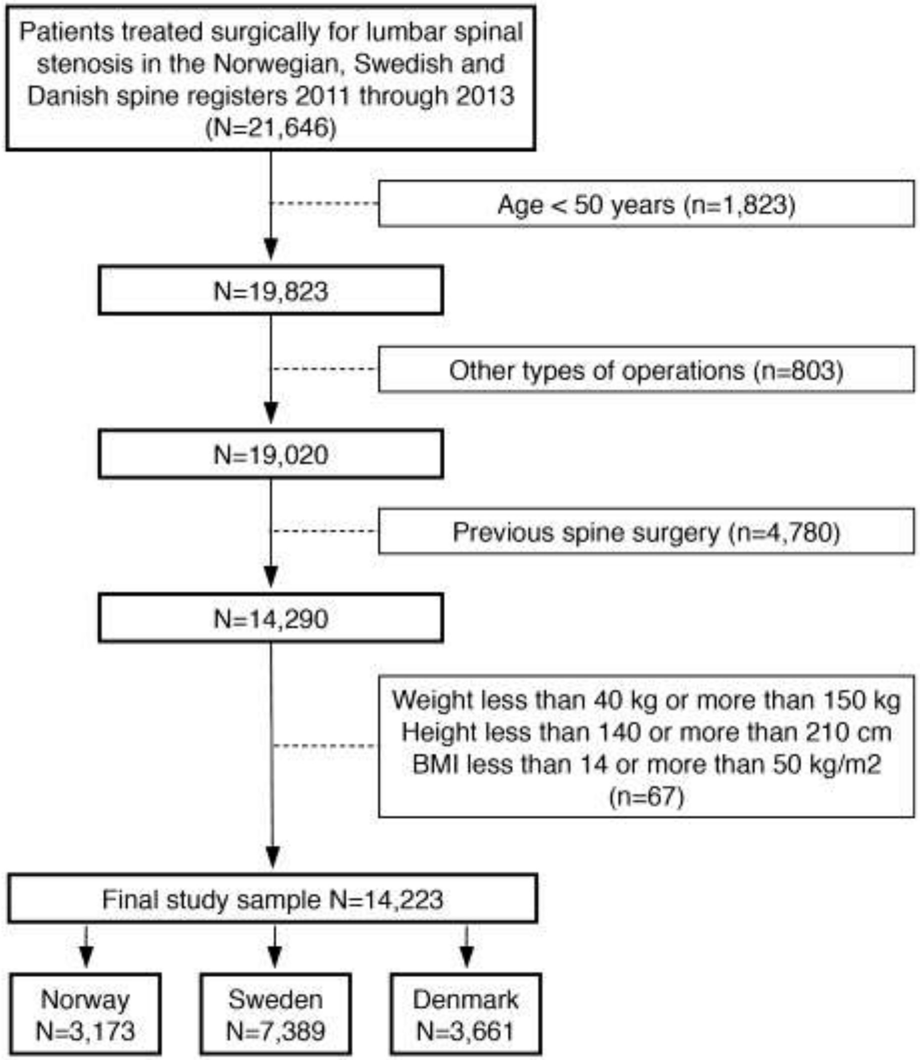
Total (n = 10890)	Norway (n =2559)	Sweden (n = 5990)	Denmark (n = 2341)	ANOVA F-test (P-value)
ODI				
ODI, mean diff (95%CI)	18 (17 to 18)	17 (17 to 18)	18 (17 to 19)	0.081*
ODI, mean diff case-mix adj (95%CI)	16 (16 to 17)	18 (18 to 19)	17 (17 to 17)	0.010*
ODI % > MCIC (95%CI)	64 (62 to 66)	60 (59 to 61)	65 (63 to 68)	<0.001‡
ODI % ≤ 22 (95%CI)	64 (62 to 66)	64 (63 to 65)	64 (61 to 66)	0.837‡
NRS leg pain				
NRS leg pain, mean diff (95%CI)	3.2 (3.1 to 3.4)	3.1 (3.0 to 3.2)	2.9 (2.7 to 3.1)	0.008*
NRS leg pain, mean diff case-mix adj (95%CI)	3.2 (3.2 to 3.3)	3.2 (3.1 to 3.2)	2.7 (2.6 to 2.8)	<0.001*
NRS leg pain % > MCIC (95%CI)	64 (62 to 66)	63 (62 to 65)	66 (63 to 68)	0.263‡
NRS back pain				
NRS back pain, mean diff (95%CI)	3.0 (2.9 to 3.2)	2.4 (2.3 to 2.4)	2.2 (2.0 to 2.3)	<0.001*
NRS back pain, mean diff case-mix adj (95%CI)	2.9 (2.9 to 3.0)	2.5 (2.4 to 2.5)	2.0 (2.0 to 2.1)	<0.001*
NRS back pain % > MCIC (95%CI)	64 (62 to 66)	61 (59 to 62)	62 (60 to 64)	0.038‡
EQ-5D				
EQ-5D, mean diff (95%CI)	0.28 (0.26 to 0.29)	0.27 (0.26 to 0.28)	0.28 (0.26 to 0.29)	0.323*
EQ-5D, mean diff case-mix adj (95%CI)	0.27 (0.26 to 0.28)	0.28 (0.28 to 0.29)	0.25 (0.24 to 0.26)	<0.001*

*Data are shown as mean (95%CI), or % (95%CI)). P-values for comparison between the groups are shown. . PROMs = Patient Reported Outcome Measures SD = Standard deviation, CI= Confidence interval, ODI = Oswestry Disability Index, NRS = Numeric Rating Scale, EQ-5D = Euro-Qol-5D 3 levels, MCIC = Minimal Clinically Important Change. Adjusted mean presented as predicted value. *Anova F-test, † Student's t test, ‡Pearson's Chi-square test*

Table 4: Improvements in PROMs by type of surgery in the combined cohort of Norway, Sweden and Denmark.

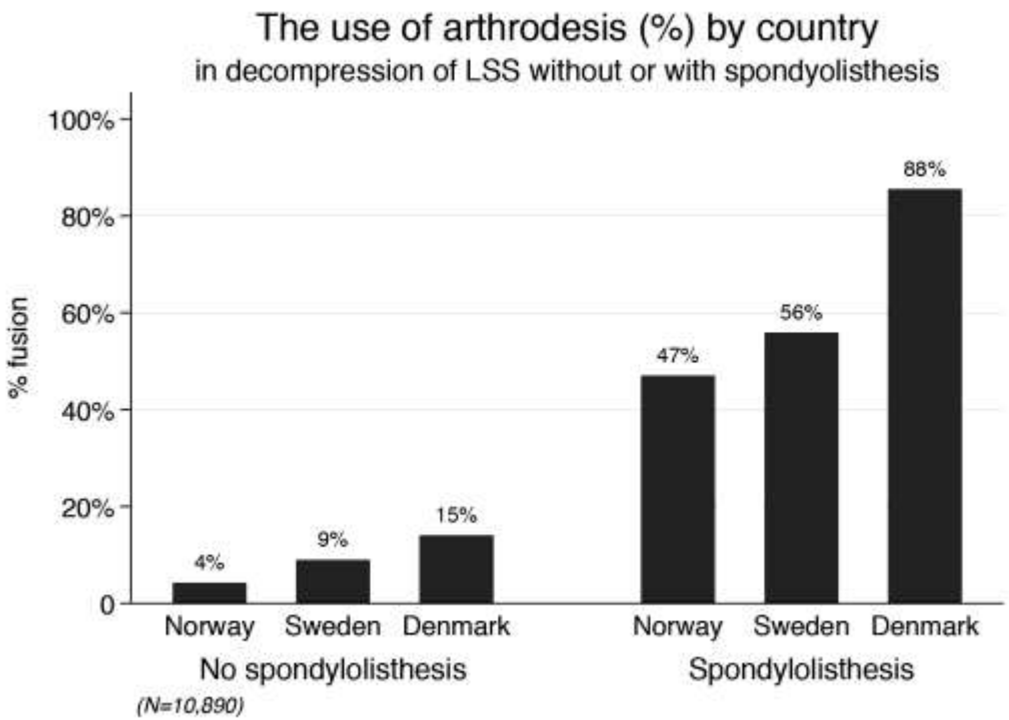
	Mean outcome difference (95%CI)			
Crude	ODI	NRS Leg Pain	NRS Back Pain	EQ-5D
No spondylolisthesis				
Decompression only (n=7791)	17 (17 to 18)	3.0 (2.9 to 3.0)	2.3 (2.3 to 2.4)	0.26 (0.25 to 0.27)
Decompression and arthrodesis (n=761)	18 (17 to 20)	3.0 (2.8 to 3.3)	2.7 (2.3 to 3.0)	0.29 (0.26 to 0.31)
p-value (Student's t-test)	*0.169*	*0.529*	*0.002*	*0.073*
Spondylolisthesis				
Decompression only (n=925)	17 (16 to 18)	3.3 (3.1 to 3.5)	2.6 (2.3 to 2.8)	0.26 (0.24 to 0.29)
Decompression and arthrodesis (n=1413)	20 (19 to 21)	3.6 (3.4 to 3.8)	3.1 (3.0 to 3.3)	0.32 (0.31 to 0.34)
p-value (Student's t-test)	*<0.001*	*0.029*	*<0.001*	*<0.001*
Case mix adjusted*				
No spondylolisthesis				
Decompression only (n=7791)	17 (17 to 18)	3.1 (3.0 to 3.1)	2.4 (2.4 to 2.5)	0.27 (0.26 to 0.27)
Decompression and arthrodesis (n=761)	19 (18 to 20)	3.0 (2.9 to 3.2)	2.7 (2.6 to 2.9)	0.30 (0.29 to 0.32)
p-value (Student's t-test)	*<0.001*	*0.710*	*<0.001*	*<0.001*
Spondylolisthesis				
Decompression only (n=925)	17 (17 to 18)	3.2 (3.1 to 3.3)	2.5 (2.4 to 2.6)	0.27 (0.25 to 0.28)
Decompression and arthrodesis (n=1413)	18 (18 to 19)	3.2 (3.1 to 3.3)	2.6 (2.5 to 2.7)	0.29 (0.28 to 0.30)
p-value (Student's t-test)	*0.010*	*0.965*	*0.077*	*0.035*

**Data adjusted for age, gender, BMI, smoking, comorbidity, and baseline PROM values. PROMs = Patient Reported Outcome Measures, ODI = Oswestry Disability Index, NRS = Numeric Rating Scale, EQ-5D = Euro-Qol-5D 3 levels.*


Patients treated surgically for lumbar spinal stenosis in the Norwegian, Swedish and Danish spine registers 2011 through 2013 (N=21,646)
Age < 50 years (n=1,823)
N=19,823
Other types of operations (n=803)
N=19,020
Previous spine surgery (n=4,780)
N=14,290
Weight less than 40 kg or more than 150 kg
Height less than 140 or more than 210 cm
BMI less than 14 or more than 50 kg/m2
(n=67)
Final study sample N=14,223
Norway N=3,173
Sweden N=7,389
Denmark N=3,661

1

Figure 1 TSJ.tif



1
2 Figure 2 TSJ.tif
3

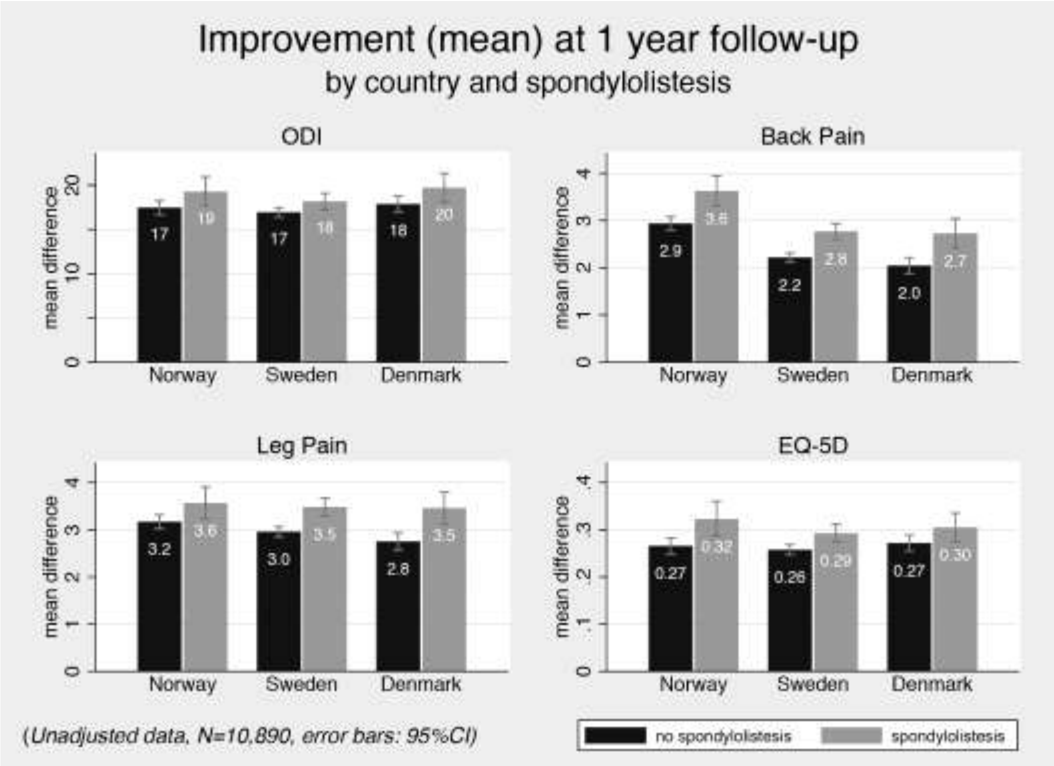


Figure 3 TSJ.tif