


Is the duration of pre-operative conservative treatment associated with the clinical outcome following surgical decompression for lumbar spinal stenosis? A study based on the Spine Tango Registry

Thomas Zweig^{1,2}  · Juliane Enke² · Anne F. Mannion³ · Rolf Sobottke⁶ · Markus Melloh⁵ · Brian J. C. Freeman⁴ · Emin Aghayev² · Spine Tango Contributors

Received: 20 February 2016/Revised: 20 October 2016/Accepted: 11 November 2016/Published online: 15 December 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract

Background The incidence of lumbar spinal stenosis (LSS) continues to rise, with both conservative and surgical management representing options for its treatment. The timing of surgery for LSS varies from shortly after the onset of symptoms to several months or years after conservative treatment. The aim of this study was to investigate the association between the duration of pre-operative conservative treatment and the ultimate outcome following surgical interventions for LSS.

Methods The study was based on prospective multicentre registry data (Spine Tango). Cases of LSS with a documented duration of conservative treatment, undergoing spinal decompression with at least one post-operative patient

assessment between 3 and 30 months, were included in the study. Cases of LSS with spondylolisthesis, additional spinal pathology or previous spinal surgery were excluded. Interrogation of the Spine Tango Registry listed 3478 patients meeting the prescribed inclusion criteria. This cohort was stratified into four groups: (1) no previous treatment ($n = 497$; 14.3%), (2) conservative treatment <6 months ($n = 965$; 27.8%), (3) conservative treatment between 6 and 12 months ($n = 758$; 21.8%), and (4) conservative treatment >12 months ($n = 1258$; 36.1%). Group 4 reference group in regression analysis. The inverse probability of treatment weighting (IPTW) was applied using the propensity score to balance the groups for their characteristics. Outcome measures included achievement of the minimum clinically important change (MCIC) score of 2 points for (a) back pain, (b) leg pain and (c) Core Outcome Measures Index (COMI), and (d) surgical complications, (e) general complications and (f) operation time >2 h.

Results Patient group (“duration of conservative therapy”) was not associated with achievement of the MCIC for post-operative relief of leg pain ($p = 0.22$), achievement of MCIC for the COMI score ($p = 0.054$), surgical complications ($p = 0.11$) or general complications ($p = 0.14$). Only MCIC for post-operative relief of back pain ($p = 0.021$) and operation time were significantly associated with patient group ($p = 0.038$). However, compared with the reference group of >12 months of conservative treatment there was no significant difference in the likelihood of achieving the MCIC for those with none, <6 or 6–12 months of conservative treatment.

Conclusions The duration of pre-operative conservative treatment was not associated with the ultimate outcome of decompression surgery. Further research is required to investigate optimal thresholds/indications for surgery and its appropriate timing in individual patients.

✉ Thomas Zweig
zweig@memcenter.unibe.ch

- ¹ The Spine Center Bern, Schänzlistrasse 39, 3000 Bern, Switzerland
- ² Swiss Center of Excellence in Medical Registries and Data Linkage (Swiss RDL), Institute for Social and Preventive Medicine, University of Bern, Finkenhubelweg 11, 3012 Bern, Switzerland
- ³ Spine Center, Department of Teaching, Research and Development, Schulthess Klinik, Lengghalde 2, 8008 Zurich, Switzerland
- ⁴ Department of Spinal Surgery, Discipline of Orthopaedics and Trauma, School of Medicine, University of Adelaide, Adelaide, Australia
- ⁵ Center for Health Sciences, School of Health Professions, Zurich University of Applied Sciences, Technikumstrasse 71, 8401 Winterthur, Switzerland
- ⁶ Centre for Orthopaedic and Trauma Surgery, Medical Center Städte Region Aachen, Mauerfeldchen 25, 52146 Würselen, Germany

Keywords Lumbar spinal stenosis · Conservative treatment · Timing of surgery · Spine Tango · Registry

Introduction

Lumbar spinal stenosis (LSS) is one of the most common degenerative spinal pathologies requiring surgical treatment in the elderly population [1]. The typical clinical presentation of LSS is neurogenic claudication resulting in leg and/or buttock pain and limited walking distance. For those over 65 years of age, the estimated incidence of LSS is reported to be 1:1000 [2, 3]. The number of patients diagnosed with LSS in Germany doubled between 2005 and 2011 from 28,000 to nearly 56,000 [4], and the continuous aging of the population can be expected to result in a further increase in its prevalence. Accordingly, the healthcare costs associated with LSS can be expected to have major economic consequences [2, 5, 6]. Despite the increased risk of surgical complications in elderly patients with comorbidities, there appears to be a disproportional rise in the frequency of lumbar spine surgery in this population [4, 6, 7].

Initial treatment of LSS is conservative, and includes physiotherapy, analgesics and targeted epidural steroid injections. If unsuccessful, surgery is advocated and has superior results when compared with continued non-operative treatment [3, 8–10]. Surgery itself, with considerable variations of techniques, provides substantial relief of pain and disability [11]. The optimal timing for surgical decompression has not yet been established. Currently, lumbar decompressive surgery may be offered to patients soon after the onset of symptoms or several months or even years later, due to the fluctuating natural history of LSS. Symptomatic LSS is mostly a chronic intermittent condition [9] and there may be a sudden exacerbation of symptoms. Both leg/buttock pain and the reduction in walking distance contribute to the burden of the condition. The fluctuating symptoms and long periods of stability make it difficult to determine the most appropriate timing for decompressive surgery. Amundsen et al. performed an analysis of the 10-year outcomes in a cohort of 100 LSS patients, of whom 31 patients had been randomly assigned to either conservative or surgical treatment [9]. Whilst mindful of the small sample size, the authors concluded that surgery for LSS seemed to be equally beneficial whether it was performed early or late (up to 3 years) after severe symptoms [9]. A recent analysis of data from the Spine Patient Outcomes Research Trial (SPORT) revealed significantly less improvement in bodily pain (SF-36) and some other outcomes in LSS patients with >12 months' duration of symptoms compared with those with a shorter

duration of symptoms, although there was no difference in post-operative leg pain between the two groups [12]. The duration of conservative treatment—as opposed to the duration of symptoms—is a rather heterogeneous variate in clinical studies, and is only rarely analyzed. More often a cut-off is used as a criterion for inclusion.

Whether an association exists between the duration of conservative treatment prior to surgery and patient outcomes is unknown. Randomized controlled trials (RCT) would be the ideal approach to further investigate the issue; however, such studies are not without their difficulties. As there is evidence for the superiority of surgery vs. conservative treatment for LSS [8–10], it may be considered unethical to randomize patients into early and delayed surgery. Furthermore, the notoriously high cross-over rates in RCTs of surgery vs. conservative treatment may negate the benefits of randomisation. In the SPORT, 46% of patients assigned to surgery and 54% of patients assigned to conservative treatment opted to change group. This considerably limits the interpretation of the results [13].

The alternatives to RCTs, observational studies (including registry studies), are not able to establish a cause-effect relationship, as it is very difficult to capture factors that influence surgeons' decisions to operate. Nonetheless, the large quantity of “real-life” data [14] captured in registries may still provide a valuable snap-shot of the outcome in patients with differing durations of conservative treatment.

The aim of this study was to investigate the association between the duration of pre-operative conservative treatment and patient- and physician-reported outcome measures observed following surgical interventions for LSS. We hypothesized that a shorter duration of conservative treatment would be associated with more favorable surgical outcomes.

Materials and methods

Spine Tango Registry

This study was carried out using the Spine Tango data pool, and is written in accordance with the STROBE statement [15]. Spine Tango, the international spine registry of EUROSPINE, the Spine Society of Europe is hosted at the University of Bern's Institute for Evaluative Research in Medicine [16]. Within the registry, patient and physician-based data are collected in a prospective observational multi-centre manner.

The last three iterations of the Spine Tango surgery form (2005, 2006, and 2011) were used in the analysis. The surgery form is a physician-based case report form

capturing information on demographics, clinical information and data on the surgical intervention. Patient-reported Core Outcome Measures Index (COMI) data are collected pre-operative and at standardized follow-up intervals (3 months, 1 and 2 years following surgery) and entered into the registry. The COMI is a short, self-administered outcome measure instrument, consisting of seven questions to evaluate the five domains: pain (back and leg/buttock pain measured separately), back-related function, symptom-specific well-being, general quality of life, and disability (social and work) [17]. Two pain Graphic Rating Scales (GRS) from 0 to 10 points are used to assess (1) back and (2) leg/buttock pain, and all other items use a five-point scale. For the summary score, the average of the scores of all five domains (each transformed to 0–10) is calculated [17]. COMI is a validated outcome tool existing in numerous languages [17–26].

Inclusion criteria and patient sample

The Spine Tango registry was examined in March 2015 for all patients with the diagnosis of degenerative LSS without degenerative spondylolisthesis. The inclusion criteria were: spinal stenosis, lumbar or lumbosacral, patient age >20 years, most severely affected segment between L1/2 and L5/S1, no previous spinal surgery, either laminotomy, hemi-laminectomy, laminectomy, and/or partial facet joint resection as type of decompression [27], known American Society of Anaesthesiologists (ASA) classification, pre-operative and at least one post-operative COMI assessment available between 3 and 30 months (Table 1). The exclusion criteria were ‘additional spinal pathology’ such as spondylolisthesis, deformity, fracture, tumor, inflammation, infection, and

reoperation. If multiple follow-up forms were available for a patient within the given follow-up period, the latest form was selected for analysis. Data from Finland, India, Moldova, The Netherlands, Singapore, Slovenia, Taiwan, and Turkey were not considered due to the lack of a validated language version of the COMI (5.4%, Table 1). The inclusion/exclusion criteria resulted in 3478 patients from 32 institutions from nine countries (Australia, Austria, Belgium, Germany, Italy, Poland, Switzerland, UK and USA).

Stratification of Cohort into four groups based on the duration of conservative treatment

Based on the duration of previous conservative treatment four groups were defined: (1) no previous treatment ($n = 497$; 14.3%), (2) conservative treatment <6 months ($n = 965$; 27.8%), (3) conservative treatment 6–12 months ($n = 758$; 21.8%), and (4) conservative treatment >12 months ($n = 1258$; 36.1%).

Outcome measures

As clinical outcomes, dichotomized achievement of the minimum clinically important change (MCIC) of two points in (a) back pain, (b) leg pain and (c) COMI score were considered. Further dichotomized outcomes were (d) surgical complications (e) general complications, and (f) operation time >2 h (Fig. 1).

Statistical analyses

The inverse probability of treatment weighting (IPTW) was applied using the propensity score to balance the

Table 1 Patient selection algorithm

Inclusion criteria	Surgery forms (January 04–March 15) $N = 76,565$	
	Included	Excluded (%)
Hospitals with a valid COMI form in the national language	72,447 (94.6%)	5.4
First dated surgery form (one per patient)	64,559 (89.1%)	10.9
Lumbar and lumbo-sacral (L1/L2–L5/S1)	46,560 (72.1%)	27.9
Spinal stenosis	17,176 (36.9%)	63.1
No additional pathology	12,681 (73.8%)	26.2
No previous surgical treatment	9641 (76.0%)	24.0
Decompression ^a	8334 (86.4%)	13.6
ASA classification known	7385 (88.6%)	11.4
Eligible for ≥ 3 months follow-up	7324 (99.2%)	0.8
Patient form (COMI) at baseline and at follow-up (3–30 months)	<u>3478</u> (47.5%)	52.5

Underline denotes endresults, i.e. total number of included patients

^a Laminotomy, hemi laminectomy, laminectomy, and/or partial facet joint resection as the type of decompression

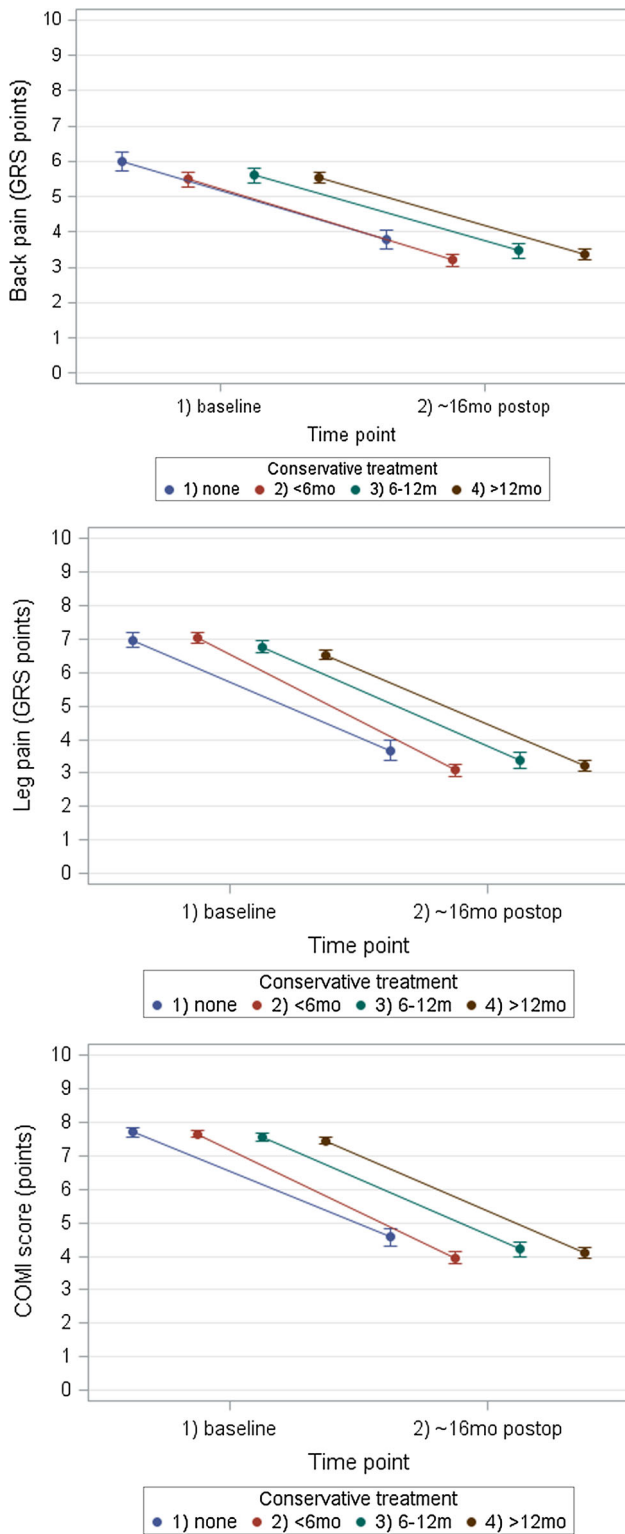


Fig. 1 Observed mean back pain relief, leg pain relief and COMI score improvement for each of the four groups of conservative treatment with 95% confidence intervals. *mo* months

groups for their baseline characteristics. An individual’s propensity score is defined by the conditional probability of being in one group or another, given the observed baseline covariates (such as age, gender, ASA classification, etc.). IPTW uses weights based on the propensity score to create a synthetic sample in which the distribution of measured baseline covariates is independent of the group assignment. Four patients with the same propensity score have an equal estimated probability of receiving no pre-operative conservative therapy, <6, 6–12, or >12 months conservative therapy. If the four patients receive four different durations of conservative therapy, the exposure allocation can be considered random, in relation to the observed covariates. Therefore, as in randomized controlled trials (RCTs), there is balance of the covariates between exposure groups after adjusting for the propensity score. The important difference between propensity score adjustment and RCTs is that the latter are able to balance both measured and unmeasured covariates. Propensity scores can only control for the measured covariates.

The following covariates were included in the propensity score: patient age (continuous), gender (male/female), ASA classification (1, 2, >2), extent of lesion (1, 2–3, >3 segments), level of surgery (L1/2–L2/3, L3/4, L4/5, L5/S1), pain reduction as the surgical goal (yes/no), functional improvement as the surgical goal (yes/no), neurological improvement as the surgical goal (yes/no), patient-reported main problem (back pain, leg pain, sensory disturbances, other), flavectomy (yes/no), partial facet joint resection (yes/no), complete facet joint resection (yes/no), laminotomy (yes/no), hemi-laminectomy (yes/no), laminectomy (yes/no), foraminotomy (yes/no), discectomy (yes/no), sequestrectomy (yes/no), other decompression (yes/no), duration of follow-up (continuous), pre-operative back pain (continuous) and leg pain (continuous), and pre-operative COMI score (continuous).

Bivariate comparison of patient characteristics in the treatment groups (before and after weighting adjustment) was performed using general linear modelling for continuous covariates and the Chi-square test for categorical covariates. Bonferroni adjustment was used for multiple testing as appropriate.

In the final step, logistic regression analysis for each outcome measure was performed and adjusted for the inverse probability of treatment weight. The covariates that remained significantly different after weighting adjustment in the bivariate comparisons were included in

the final models. For the calculation of odds ratios, conservative treatment duration of >12 months was used as the reference group, as it was the largest treatment group. A post hoc analysis was also performed with conservative treatment duration of <6 months as the reference group (the group that showed the greatest difference from the other groups) to identify the location of the specific group differences revealed by the significant p value for the overall association (Table 3 in “Appendix”).

The alpha level was set at 0.05 throughout the study. All statistical analyses were conducted using SAS 9.4 (SAS Institute, Inc., Cary, NC, USA).

Results

Before weighting adjustment, the groups were different with respect to the independent variables patient age, percentage with disc herniation, ASA classification, extent of lesion, level of surgery, goal of treatment, percentages undergoing flavectomy, laminectomy, hemi-laminectomy, sequestrectomy, partial facet joint resection, other decompression type, mean follow-up interval, and back pain, leg pain, and COMI score at baseline (Table 2). The non-adjusted (before weighting) bivariate comparisons also showed significant differences between the groups for the outcomes post-operative back pain, leg pain, leg pain relief, COMI score, change in COMI score, percentages achieving MCIC for leg pain and for COMI, percentage of surgical complications, and percentage of those with operation time >2 h (Table 2).

The weighting adjustment was successful in balancing the patient group for all patient characteristics ($p > 0.45$) except for the proportions of patients undergoing sequestrectomy ($p = 0.006$) and foraminotomy ($p = 0.029$) (Table 2).

After weighting adjustment, differences in the following outcomes were observed: back pain relief (maximum difference in means 0.4; $p = 0.047$); proportions achieving MCIC in back pain (maximum difference in proportions 6.8%; $p = 0.015$); COMI score change (maximum difference in means 0.3; $p = 0.042$); and proportion with operation time >2 h (maximum difference in proportions 5.9%; $p = 0.040$) all in favour of group 2 (Table 2).

Logistic regression analyses

As the proportions of patients undergoing sequestrectomy and foraminotomy remained significantly different after

weighting adjustment (see above, and Table 2), both covariates were included in the final models.

Patient group (i.e., duration of conservative treatment) was not associated with achievement of the MCIC in leg pain at follow-up ($p = 0.22$).

Overall, patient group was associated with achievement of the MCIC in back pain ($p = 0.021$). However, compared with the reference group of >12 months of conservative treatment, there was no significant difference in the likelihood of achieving the MCIC for those with none, <6 or 6–12 months of conservative treatment (Fig. 2). In the post hoc analysis using conservative treatment duration of <6 months as the reference group, the odds of achieving the MCIC for back pain were 0.77-times lower (95% CI 0.62–0.96) for the groups with no conservative treatment or treatment between 6 and 12 months (95% CI 0.63–0.93) (Table 3 in “Appendix”).

The duration of conservative treatment showed no significant association with achievement of the MCIC for the COMI ($p = 0.054$; Fig. 2).

The duration of conservative treatment was not associated with the proportion of surgical or general complications ($p = 0.11$ and $p = 0.14$, respectively).

Overall, the duration of conservative treatment was associated with the duration of the operation ($p = 0.038$). However, compared with the reference group of >12 months of conservative treatment, there was no significant difference in the duration of the operation for those with none, <6 or 6–12 months of conservative treatment (Fig. 2). In the post hoc analysis, compared with patients <6 months’ conservative treatment (reference group), the odds of the operation lasting more than 2 h were 1.39 (95% CI 1.11–1.74) higher for the group with 6–12 months conservative treatment (Table 3 in “Appendix”).

Discussion

We hypothesized that a short duration of conservative treatment would be associated with more favorable surgical outcomes. However, the duration of conservative treatment prior to surgery showed no significant association with the achievement of a clinically important relief of leg pain or improvement in the COMI score. Furthermore, there was no association between duration of conservative treatment and surgical or general complications. Only clinically important back pain relief showed significant differences between the groups: it was seen more frequently in patients with a short (<6 months) duration of conservative treatment compared with those with no conservative treatment

Table 2 Characteristics of patients and outcomes before and after propensity score based weighting adjustment in groups

Patient characteristics	Before weighting				After weighting				Comparison (p value)
	(1) No previous treatment	(2) <6 months	(3) 6–12 months	(4) >12 months	(1) No previous treatment	(2) <6 months	(3) 6–12 months	(4) >12 months	
N (%)	497 (14.3)	965 (27.8)	758 (21.8)	1258 (36.1)	497 (14.3)	965 (27.8)	758 (21.8)	1258 (36.1)	–
Mean age ± SD (years)	67.6 ± 12.1	66.3 ± 12.5	67.4 ± 11.4	69.0 ± 10.9	68.1 ± 12.1	67.9 ± 11.9	67.9 ± 11.9	67.6 ± 11.2	0.86
Females (%)	44.1	48.0	43.5	48.3	47.3	47.1	46.4	47.1	0.99
Disc herniation (%)	24.6	26.3	23.5	18.0	23.8	21.1	22.8	23.1	0.59
ASA 1 (%)	16.9	21.6	17.0	13.5	14.9	16.5	16.4	16.4	0.56
ASA 2 (%)	58.0	59.8	61.5	56.8	55.5	57.8	58.3	58.7	
ASA >2 (%)	25.2	18.7	21.5	29.7	29.7	25.7	25.3	24.9	
Extent of lesion: monosegmental (%)	41.7	49.4	41.2	30.1	35.9	38.3	38.9	38.8	0.72
Extent of lesion: bi-/trisegmental (%)	43.9	46.3	51.5	57.2	55.1	51.0	51.6	51.9	
Extent of lesion: >3 segments (%)	14.5	4.3	7.4	12.7	9.0	10.7	9.5	9.3	
L1/2–L2/3 (%)	6.2	6.9	7.1	6.3	5.3	6.5	6.8	7.0	0.93
L3/4 (%)	28.0	21.7	22.7	25.3	21.9	23.6	23.8	24.0	
L4/5 (%)	51.9	57.2	59.1	57.9	60.8	58.2	57.1	56.7	
L5/S1 (%)	13.9	14.2	11.1	10.6	12.1	11.7	12.2	12.3	
Treatment goal: pain relief (%)	95.2	96.2	95.9	96.3	96.0	95.9	95.7	96.2	0.97
Treatment goal: functional improvement (%)	57.6	64.3	65.2	69.2	66.1	64.7	64.4	65.4	0.91
Treatment goal: neurological improvement (%)	22.7	27.7	19.8	22.7	22.4	21.6	23.1	24.4	0.46
Main problem: back pain (%)	26.6	22.8	26.5	28.5	26.6	26.7	27.2	26.6	1.0
Main problem: leg pain (%)	46.1	50.4	48.6	48.6	49.5	48.6	47.9	48.8	
Main problem: sensory disturbances (%)	27.0	25.9	24.0	22.4	23.1	23.8	24.2	23.8	
Main problem: other (%)	0.4	0.9	0.9	0.9	0.7	0.9	0.8	0.9	
Flavectomy (%)	61.2	71.7	75.3	82.4	75.8	74.7	74.8	74.7	0.96
Facet joint resection partial (%)	53.3	67.3	68.7	76.4	67.6	68.5	69.3	68.8	0.93
Laminotomy (%)	51.9	46.0	55.8	63.8	57.7	57.3	56.0	55.9	0.85
Foraminotomy (%)	32.4	41.1	50.5	48.0	52.3	45.0	44.8	45.7	0.029
Laminectomy (%)	29.2	17.2	19.3	11.1	15.3	17.6	17.1	17.3	0.72
Discectomy (%)	22.1	23.3	21.4	22.7	25.0	22.0	22.9	22.4	0.60
Hemi-laminectomy (%)	12.7	15.9	12.4	10.1	12.1	12.4	13.2	14.0	0.63

Table 2 continued

Patient characteristics	Before weighting				After weighting				
	(1) No previous treatment	(2) <6 months	(3) 6–12 months	(4) >12 months	(1) No previous treatment	(2) <6 months	(3) 6–12 months	(4) >12 months	Comparison (p value)
Sequestrectomy (%) ^a	5.8	12.6	8.8	8.4	14.5	9.1	9.5	9.7	0.006
Facet joint resection full (%)	1.6	2.2	2.2	3.6	2.4	2.6	2.7	2.8	0.98
Other decompression (%)	2.0	4.3	4.6	2.3	3.6	3.1	3.3	3.5	0.95
Follow-up interval \pm SD (months)	14.7 \pm 8.4	15.2 \pm 8.5	16.1 \pm 8.5	16.9 \pm 8.4	16.5 \pm 8.6	16.3 \pm 8.6	16.0 \pm 8.4	16.0 \pm 8.5	0.53
Back pain at baseline \pm SD (points)	6.0 \pm 3.0	5.5 \pm 3.1	5.6 \pm 3.0	5.5 \pm 2.9	5.4 \pm 3.2	5.6 \pm 3.0	5.6 \pm 3.0	5.6 \pm 2.9	0.45
Leg pain at baseline \pm SD (points)	7.0 \pm 2.5	7.0 \pm 2.4	6.8 \pm 2.5	6.5 \pm 2.5	6.7 \pm 2.6	6.8 \pm 2.6	6.7 \pm 2.5	6.7 \pm 2.5	0.92
COMI score at baseline \pm SD (points)	7.7 \pm 1.6	7.7 \pm 1.8	7.6 \pm 1.8	7.4 \pm 1.8	7.5 \pm 1.7	7.6 \pm 1.8	7.5 \pm 1.8	7.5 \pm 1.7	0.58
Outcomes									
Back pain postoperative \pm SD (points)	3.8 \pm 3.1	3.2 \pm 2.8	3.5 \pm 3.0	3.4 \pm 2.8	3.3 \pm 3.0	3.1 \pm 2.8	3.4 \pm 3.0	3.4 \pm 2.9	0.14
Back pain relief \pm SD (points)	2.2 \pm 3.5	2.3 \pm 3.4	2.1 \pm 3.2	2.2 \pm 3.2	2.1 \pm 3.5	2.5 \pm 3.3	2.1 \pm 3.2	2.2 \pm 3.2	0.047
Achieved MCIC in back pain (%)	52.1	56.2	53.0	57.2	52.5	59.3	52.8	56.8	0.015
Leg pain postoperative \pm SD (points)	3.7 \pm 3.4	3.1 \pm 3.0	3.4 \pm 3.1	3.2 \pm 3.0	3.2 \pm 3.3	3.0 \pm 3.0	3.4 \pm 3.1	3.2 \pm 3.1	0.14
Leg pain relief \pm SD (points)	3.3 \pm 3.8	3.9 \pm 3.5	3.4 \pm 3.4	3.3 \pm 3.6	3.5 \pm 3.7	3.7 \pm 3.6	3.4 \pm 3.4	3.5 \pm 3.6	0.12
Achieved MCIC in leg pain (%)	65.2	72.4	67.6	68.4	68.1	71.1	67.3	70.4	0.27
COMI score postoperative \pm SD (points)	4.6 \pm 3.0	4.0 \pm 2.9	4.2 \pm 2.9	4.1 \pm 2.8	4.1 \pm 3.0	3.9 \pm 2.9	4.2 \pm 2.9	4.1 \pm 2.9	0.17
COMI score change \pm SD (points)	3.1 \pm 2.9	3.7 \pm 3.0	3.4 \pm 2.8	3.3 \pm 2.9	3.4 \pm 2.9	3.7 \pm 3.0	3.4 \pm 2.8	3.4 \pm 2.9	0.042
Achieved MCIC in COMI score (%)	60.4	68.5	65.6	65.4	65.4	70.3	65.6	65.6	0.06
Surgical complication (%)	11.7	9.4	7.3	11.1	11.3	10.1	7.8	10.9	0.12
Dura lesion (%)	8.7	7.5	5.4	7.9	8.7	8.2	5.9	8.1	0.17
General complication (%)	4.4	4.1	3.3	5.1	5.9	5.3	3.4	4.4	0.13

Table 2 continued

Patient characteristics	Before weighting				After weighting				
	(1) No previous treatment	(2) <6 months	(3) 6–12 months	(4) >12 months	(1) No previous treatment	(2) <6 months	(3) 6–12 months	(4) >12 months	Comparison (p value)
Operation time >2 h (%)	17.9	15.1	25	27.6	22.7	20.7	26.6	23.4	0.040

SD standard deviation, MCIC clinically important change

^a Overall, sequestrectomy was observed in 9.3% of patients, and out of ten possible types of decompression it was ranked 8th. It is possible that herniated disc occurred together with stenosis (at the same or different level) and this might account for its occasional endorsement on the form

or treatment lasting between 6 and 12 months but not compared with those with more than 12 months’ conservative treatment. Nonetheless, the clinical relevance of the statistically significant differences was questionable (see later).

Clinical implications

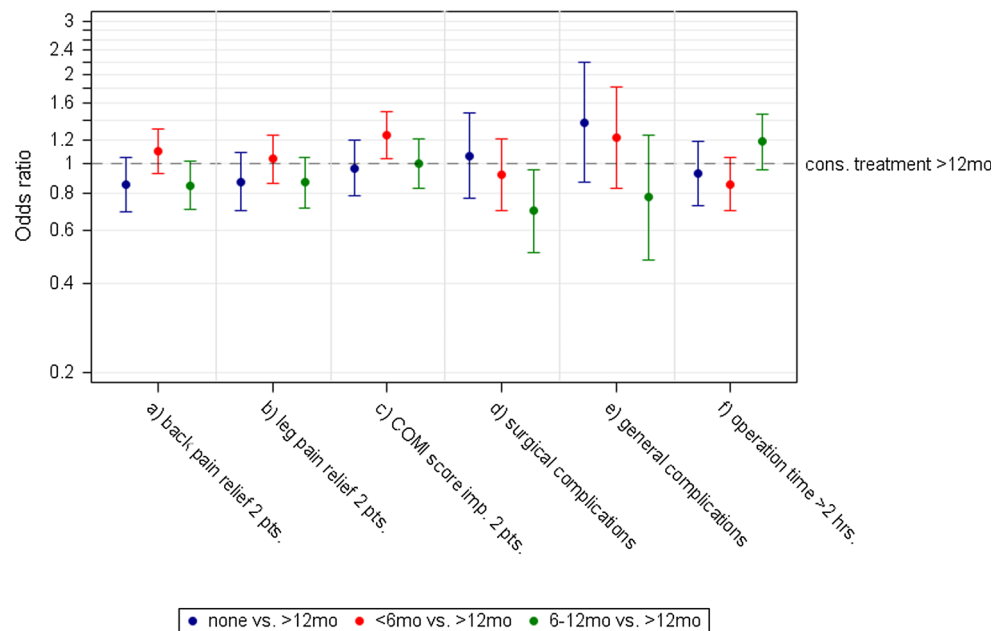
This study has shown the mere duration of pre-operative conservative treatment was, by and large, not associated with clinical outcome following surgical decompression for lumbar spinal stenosis.

Radcliff et al. found some outcomes were significantly better for patients with pre-operative symptoms for less than 12 months, although there was no difference between the two groups for post-operative leg pain [12]. They observed significantly less improvement in bodily pain (SF-36) in patients with >12 months’ duration of symptoms vs. a shorter duration of symptoms [12]. In the present study, the maximum difference in the proportion of patients achieving clinically relevant back pain relief in the different groups was 7%. This difference is potentially of low clinical relevance, as the corresponding number needed to treat (NNT) is 15 [28]. Furthermore, the primary symptom in LSS is leg and/or buttock pain, not back pain.

In our series, the proportion of long (>2 h) operations was lower for those with pre-operative conservative treatment less than 6 months than for those patients with 6–12 months’ conservative treatment (but not different compared with the other durations of conservative treatment). However, as with back pain relief, the size of the effect may render it of limited clinical relevance. Radcliff et al. did not identify a significant relationship between duration of symptoms (≤12 or >12 months) and duration of operation in a bivariate comparison [12].

In our study, conservative treatment >12 months was set as the reference group, because the greatest number of patients fell into this category and it had been used as a cut-off for forming sub-groups in a previous study [12]. However, even where a significant overall effect of duration of conservative treatment was observed (for MCIC for back pain and duration of operation), the differences between the reference group (conservative treatment >12 months) and the other treatment groups were not significant for any of the outcomes. A post hoc analysis using <6 months’ treatment as the reference group revealed that there was a significantly higher proportion of patients with clinically important back pain relief in this group compared with no conservative treatment or a duration of between 6 and 12 months; further, conservative treatment <6 months was less likely to be

Fig. 2 Plot of odds ratios with 95% confidence intervals. For the outcomes **a–c**, the odds ratios with confidence intervals below the reference line demonstrate inferior results; for the outcomes **d–f** the odds ratios with confidence intervals below the reference line demonstrate superior results. Significant *p* values are given in *black*. *mo* months



associated with a long surgery in comparison with conservative treatment 6–12 months (though not compared with no conservative treatment or a duration of treatment >12 months). However, the small size of the differences, and their difficult interpretation, render them of unlikely clinical relevance.

LSS can be considered a chronic condition with fluctuating symptoms. Approximately one third of surgically treated LSS patients in the spine registry had had pre-operative conservative treatment for 12 months or longer. Surgery with no previous conservative treatment was performed in 14.3%. These patients had the highest back pain, leg pain and COMI scores at baseline, while the groups with a longer duration of conservative treatment had lower mean pain and COMI score levels. However, the differences between groups were at most 0.5 GRS points, and, although statistically significant, may be of limited clinical relevance.

Strengths and weaknesses

We showed that there was no clinically relevant association between the duration of conservative treatment and outcome of surgery for LSS. One must, however, be careful when interpreting these findings. The fact that there was no association most likely confirms the appropriateness of the treatment hierarchy used by the contributors to the registry, as the final results in all “duration” groups were similar. This is not to say that a person who was treated for a short

time would have done just as well if they had persevered for a longer time, or vice versa; it simply shows that the chosen length of conservative care before surgery was probably appropriate for the given patient. However, the decision-making process for surgery was not documented, and the design of the study does not enable us to draw conclusions in relation to this.

The types of conservative treatment (e.g., physiotherapy, analgesic medication, neuromodulators, and epidural steroid injections) are not explicitly documented in the registry. This is a limitation of the study. It can only be speculated that these treatments were equally distributed between groups.

Despite the multinational registry setting and the large number of participating hospitals, a follow-up rate of just 47.5% was achieved. Clearly, this is a limitation of the study. The comparison of patients with and without COMI forms/follow-ups (Table 4 in “Appendix”) did not suggest a selection bias towards patients with a likely favorable outcome. Data from nine countries were included in the analysis. Cultural and healthcare system differences may have potentially influenced the results of the study as thresholds for surgery may differ between institutions and countries. Detailed analyses of countries and individual centres, with larger sample sizes, are required to investigate possible differences. The distribution of patient groups by country is shown in Table 5 in “Appendix”.

This study used a robust methodological approach to balance patient groups for their characteristics, to adjust for

selection bias and confounding by indication in relation to the measured factors. Four main propensity score adjustment methods exist: propensity score matching, IPTW, stratification on the propensity score, and covariate adjustment using the propensity score [29]. The analysis of a propensity score-matched sample can mimic that of an RCT [29]. However, this type of adjustment is typically used for comparison of two groups. In some settings, propensity score matching and IPTW remove systematic differences between studied subject groups to a comparable degree [29]. Based on these considerations, and the fact that we had more than two groups, the IPTW was chosen.

Conclusions

Duration of pre-operative conservative treatment was not associated with differences in post-operative relief of leg pain, COMI score improvement, or surgical or general

medical complication rates. Longer conservative treatment does not appear to be detrimental to the ultimate surgical outcome following lumbar spinal decompression. Further research is required to investigate optimal thresholds/indications for surgery and its appropriate timing in individual patients.

Compliance with ethical standards

Conflict of interest None.

Appendix

See Tables 3, 4 and 5.

Table 3 Results of the models with the conservative treatment <6 months as reference group (post hoc analysis)

Outcome	Effect	Odds ratio	Lower 95% confidence interval	Upper 95% confidence interval
(a) Back pain relief 2 points	(1) None vs. (2) <6 months	0.77	0.62	0.96
	(3) 6–12 vs. (2) <6 months	0.77	0.63	0.93
	(4) >12 vs. (2) <6 months	0.91	0.76	1.07
(b) Leg pain relief 2 points	(1) None vs. (2) <6 months	0.84	0.67	1.06
	(3) 6–12 vs. (2) <6 months	0.84	0.68	1.03
	(4) >12 vs. (2) <6 months	0.96	0.80	1.16
(c) COMI score improvement 2 points	(1) None vs. (2) <6 months	0.78	0.62	0.98
	(3) 6–12 vs. (2) <6 months	0.81	0.66	0.99
	(4) >12 vs. (2) <6 months	0.80	0.67	0.96
(d) Surgical complications	(1) None vs. (2) <6 months	1.16	0.82	1.63
	(3) 6–12 vs. (2) <6 months	0.76	0.54	1.06
	(4) >12 vs. (2) <6 months	1.09	0.83	1.43
(e) General complications	(1) None vs. (2) <<6 months	1.12	0.71	1.78
	(3) 6–12 vs. (2) <6 months	0.63	0.39	1.02
	(4) >12 vs. (2) <6 months	0.81	0.55	1.20
(f) Operation time >2 h	(1) None vs. (2) <6 months	1.09	0.84	1.41
	(3) 6–12 vs. (2) <6 months	1.39	1.11	1.74
	(4) >12 vs. (2) <6 months	1.17	0.95	1.43

Significant effects are bold

Table 4 Comparison of the included patients with COMI forms and those without COMI forms at baseline and/or follow-up

Characteristics	Included patients: with COMI forms	Patients without COMI forms	Comparison (<i>p</i> value)	Overall
<i>N</i> patients (%)	3478 (47.5)	3846 (52.5)	–	7324 (100.0)
(1) No pre-operative treatment (%)	14.3	14.1	<0.001	14.2
(2) Pre-operative cons. treatment <6 months (%)	27.8	33.7		30.9
(3) Pre-operative cons. treatment 6–12 months (%)	21.8	21.5		21.6
(4) Pre-operative cons. treatment >12 months (%)	36.1	30.7		33.3
Mean age (SD) (years)	67.7 (11.7)	65.4 (13.4)	<0.001	50.2 (11.1)
Age range (years)	22–97	20–101	–	20–101
Females (%)	46.6	48.6	0.08	47.6
Disc herniation (%)	22.4	31.0	<0.001	26.9
ASA 1 (%)	17.0	21.8	<0.001	19.5
ASA 2 (%)	58.8	56.4		57.5
ASA >2 (%)	24.2	21.8		23.0
Extent of lesion: monosegmental (%)	39.5	54.2	<0.001	47.2
Extent of lesion: bi-/trisegmental (%)	51.0	41.1		45.8
Extent of lesion: >3 segments (%)	9.5	4.7		7.0
L1/2–L2/3 (%)	6.6	6.5	<0.001	6.6
L3/4 (%)	24.1	21.1		22.6
L4/5 (%)	57.1	57.5		57.3
L5/S1 (%)	12.1	14.9		13.6
Treatment goal: pain relief (%)	96.0	97.9	<0.001	97.0
Treatment goal: functional improvement (%)	65.3	66.3	0.34	65.8
Treatment goal: neurological improvement (%)	23.4	36.3	<0.001	30.2
Flavectomy (%)	74.9	65.9	<0.001	70.2
Facet joint resection partial (%)	68.9	58.1	<0.001	63.2
Laminotomy (%)	55.4	46.8	<0.001	50.9
Foraminotomy (%)	44.4	39.2	<0.001	41.7
Laminectomy (%)	17.2	27.4	<0.001	22.5
Discectomy (%)	22.5	28.1	<0.001	25.4
Hemi-laminectomy (%)	12.6	12.9	0.69	12.7
Sequestrectomy (%)	9.3	13.1	<0.001	11.3
Facet joint resection full (%)	2.6	1.3	<0.001	1.9
Other decompression (%)	3.3	6.2	<0.001	4.9

Chi-square test was used for comparison. Statistically significant differences are in bold

The table presents a comparison of patients with and without COMI forms. The distribution for the duration of pre-operative conservative treatment was significantly different between the selected and non-selected patients: there were 6% fewer patients with a short duration (<6 months) and 5% more with a long duration (>12 months) in the selected group compared with the remaining population

The vast majority of the significant differences shown in the Table were small and, therefore, not likely to have been clinically relevant. Two observed differences may support a selection bias towards simpler cases, namely the lower proportion of patients with laminectomy and with neurological improvement as the treatment goal. In contrast, several findings such as older age, higher ASA status, multi-segmental surgery, and higher proportions of decompressive measures point instead to a selection towards less favorable cases. Overall, the results of this analysis did not suggest a bias towards the selection of favorable cases only

Table 5 Distribution of the patient groups by country

Country	Patient group			
	(1) None	(2) <6 months	(3) 6–12 months	(4) >12 months
Australia	5.7 (8)	36.4 (51)	30.0 (42)	27.9 (39)
Austria	100.0 (18)	–	–	–
Belgium	23.6 (113)	25.3 (121)	15.5 (74)	35.7 (171)
Germany	2.1 (10)	42.4 (199)	16.2 (76)	39.2 (184)
Italy	100.0 (1)	–	–	–
Poland	7.7 (1)	46.2 (6)	23.1 (3)	23.1 (3)
Switzerland	7.6 (120)	23.9 (379)	23.1 (367)	45.5 (723)
UK	31.0 (224)	27.4 (198)	25.2 (182)	16.5 (119)
USA	4.4 (2)	23.9 (11)	30.4 (14)	41.3 (19)

References

- Sobottke R, Herren C, Siewe J, Mannion AF, Roder C, Aghayev E (2015) Predictors of improvement in quality of life and pain relief in lumbar spinal stenosis relative to patient age: a study based on the Spine Tango Registry. *Eur Spine J*. doi:10.1007/s00586-015-4078-8
- Joaquim AF, Sansur CA, Hamilton DK, Shaffrey CI (2009) Degenerative lumbar stenosis: update. *Arq Neuropsiquiatr* 67:553–558
- Benoist M (2002) The natural history of lumbar degenerative spinal stenosis. *Joint Bone Spine* 69:450–457
- Kalff R, Ewald C, Waschke A, Gobisch L, Hopf C (2013) Degenerative lumbar spinal stenosis in older people: current treatment options. *Deutsch Arztebl Int* 110:613–623. doi:10.3238/arztebl.2013.0613 (quiz 624)
- Katz JN, Losina E (2008) Cost-effectiveness of spine surgery: the jury is out. *Ann Intern Med* 149:901–903
- Sobottke R, Aghayev E, Roder C, Eysel P, Delank SK, Zweig T (2012) Predictors of surgical, general and follow-up complications in lumbar spinal stenosis relative to patient age as emerged from the Spine Tango Registry. *Eur Spine J* 21:411–417. doi:10.1007/s00586-011-2016-y
- Atlas SJ, Keller RB, Wu YA, Deyo RA, Singer DE (2005) Long-term outcomes of surgical and nonsurgical management of lumbar spinal stenosis: 8 to 10 year results from the maine lumbar spine study. *Spine* 30:936–943
- Malmivaara A, Slati P, Heliövaara M, Sainio P, Kinnunen H, Kankare J, Dalin-Hirvonen N, Seitsalo S, Herno A, Kortekangas P, Niinimäki T, Ronty H, Tallroth K, Turunen V, Knekt P, Harkanen T, Hurri H, Finnish Lumbar Spinal Research G (2007) Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. *Spine* 32:1–8. doi:10.1097/01.brs.0000251014.81875.6d
- Amundsen T, Weber H, Nordal HJ, Magnaes B, Abdelnoor M, Lilleas F (2000) Lumbar spinal stenosis: conservative or surgical management?: a prospective 10-year study. *Spine* 25:1424–1435 (discussion 1435–1426)
- Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Blood E, Hanscom B, Herkowitz H, Cammisa F, Albert T, Boden SD, Hilibrand A, Goldberg H, Berven S, An H, Investigators S (2008) Surgical versus nonsurgical therapy for lumbar spinal stenosis. *N Engl J Med* 358:794–810. doi:10.1056/NEJMoa0707136
- Fritsch CG, Ferreira ML, Maher CG, Herbert RD, Pinto RZ, Koes B, Ferreira PH (2016) The clinical course of pain and disability following surgery for spinal stenosis: a systematic review and meta-analysis of cohort studies. *Eur Spine J*. doi:10.1007/s00586-016-4668-0
- Radcliff KE, Rihn J, Hilibrand A, DiIorio T, Tosteson T, Lurie JD, Zhao W, Vaccaro AR, Albert TJ, Weinstein JN (2011) Does the duration of symptoms in patients with spinal stenosis and degenerative spondylolisthesis affect outcomes?: analysis of the Spine Outcomes Research Trial. *Spine* 36:2197–2210. doi:10.1097/BRS.0b013e3182341edf
- Weinstein JN, Lurie JD, Tosteson TD, Zhao W, Blood EA, Tosteson AN, Birkmeyer N, Herkowitz H, Longley M, Lenke L, Emery S, Hu SS (2009) Surgical compared with nonoperative treatment for lumbar degenerative spondylolisthesis. four-year results in the Spine Patient Outcomes Research Trial (SPORT) randomized and observational cohorts. *J Bone Jt Surg Am* 91:1295–1304. doi:10.2106/JBJS.H.00913
- Staub LP, Ryser C, Roder C, Mannion AF, Jarvik JG, Aebi M, Aghayev E (2016) Total disc arthroplasty versus anterior cervical interbody fusion: use of the Spine Tango registry to supplement the evidence from randomized control trials. *Spine J* 16:136–145. doi:10.1016/j.spinee.2015.11.056
- von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, Initiative S (2007) The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Bull World Health Organ* 85:867–872
- Melloh M, Staub L, Aghayev E, Zweig T, Barz T, Theis JC, Chavanne A, Grob D, Aebi M, Roeder C (2008) The international spine registry SPINE TANGO: status quo and first results. *Eur Spine J* 17:1201–1209. doi:10.1007/s00586-008-0665-2
- Mannion AF, Elfering A, Staerkle R, Junge A, Grob D, Semmer NK, Jacobshagen N, Dvorak J, Boos N (2005) Outcome assessment in low back pain: how low can you go? *Eur Spine J* 14:1014–1026. doi:10.1007/s00586-005-0911-9
- Mannion AF, Porchet F, Kleinstück FS, Lattig F, Jeszenszky D, Bartanusz V, Dvorak J, Grob D (2009) The quality of spine surgery from the patient's perspective. Part 1: the Core Outcome Measures Index in clinical practice. *Eur Spine J* 18(Suppl 3):367–373. doi:10.1007/s00586-009-0942-8
- Monticone M, Ferrante S, Maggioni S, Grenat G, Checchia GA, Testa M, Teli MG, Mannion AF (2014) Reliability, validity and responsiveness of the cross-culturally adapted Italian version of the core outcome measures index (COMI) for the neck. *Eur Spine J* 23:863–872. doi:10.1007/s00586-013-3092-y
- Genevay S, Cedraschi C, Marty M, Rozenberg S, De Goumoens P, Faundez A, Balague F, Porchet F, Mannion AF (2012) Reliability and validity of the cross-culturally adapted French version of the Core Outcome Measures Index (COMI) in patients with low back pain. *Eur Spine J* 21:130–137. doi:10.1007/s00586-011-1992-2
- Ferrer M, Pellise F, Escudero O, Alvarez L, Pont A, Alonso J, Deyo R (2006) Validation of a minimum outcome core set in the

- evaluation of patients with back pain. *Spine* 31:1372–1379. doi:[10.1097/01.brs.0000218477.53318.bc](https://doi.org/10.1097/01.brs.0000218477.53318.bc) (**discussion 1380**)
22. Storheim K, Brox JI, Lochting I, Werner EL, Grotle M (2012) Cross-cultural adaptation and validation of the Norwegian version of the Core Outcome Measures Index for low back pain. *Eur Spine J* 21:2539–2549. doi:[10.1007/s00586-012-2393-x](https://doi.org/10.1007/s00586-012-2393-x)
23. Damasceno LH, Rocha PA, Barbosa ES, Barros CA, Canto FT, Defino HL, Mannion AF (2012) Cross-cultural adaptation and assessment of the reliability and validity of the Core Outcome Measures Index (COMI) for the Brazilian-Portuguese language. *Eur Spine J* 21:1273–1282. doi:[10.1007/s00586-011-2100-3](https://doi.org/10.1007/s00586-011-2100-3)
24. Qiao J, Zhu F, Zhu Z, Xu L, Wang B, Yu Y, Qian BP, Ding Y, Qiu Y (2013) Validation of the simplified Chinese version of the Core Outcome Measures Index (COMI). *Eur Spine J* 22:2821–2826. doi:[10.1007/s00586-013-2761-1](https://doi.org/10.1007/s00586-013-2761-1)
25. Miekisiak G, Kollataj M, Dobrogowski J, Kloc W, Libionka W, Banach M, Latka D, Sobolewski T, Sulewski A, Nowakowski A, Kiwic G, Pala A, Potaczek T (2013) Cross-cultural adaptation and validation of the Polish version of the core outcome measures index for low back pain. *Eur Spine J* 22:995–1001. doi:[10.1007/s00586-012-2607-2](https://doi.org/10.1007/s00586-012-2607-2)
26. Klemencsics I, Lazary A, Valasek T, Szoverfi Z, Bozsodi A, Eltes P, Fekete TF, Varga PP (2016) Cross-cultural adaptation and validation of the Hungarian version of the Core Outcome Measures Index for the back (COMI Back). *Eur Spine J* 25:257–264. doi:[10.1007/s00586-014-3750-8](https://doi.org/10.1007/s00586-014-3750-8)
27. Definitions of the degenerative diagnosis groups. EUROSPINE, Spine Tango committee. http://www.eurospine.org/cm_data/def_of_degen_patho.pdf. August 2016
28. Citrome L (2008) Compelling or irrelevant? Using number needed to treat can help decide. *Acta Psychiatr Scand* 117:412–419. doi:[10.1111/j.1600-0447.2008.01194.x](https://doi.org/10.1111/j.1600-0447.2008.01194.x)
29. Austin PC (2011) An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivar Behav Res* 46:399–424. doi:[10.1080/00273171.2011.568786](https://doi.org/10.1080/00273171.2011.568786)