



Contents lists available at ScienceDirect

North American Spine Society Journal (NASSJ)

journal homepage: www.elsevier.com/locate/xnsj

Clinical Studies

One- and 2-year outcomes of lumbar facet arthroplasty versus spinal fusion in young (<65 years) and old (≥65 years) patients for the treatment of degenerative spondylolisthesis and stenosis



Annabelle Shaffer, MS^a, Alexander K. Yu, MD^b, Albert Yu, PhD^c, Graham Huesmann, MD PhD^{a,d}, Ravishanker Iyer, PhD^e, Paul M. Arnold, MD^{a,f,*}

^a Carle Illinois College of Medicine, University of Illinois Urbana-Champaign, Urbana, IL 61801, United States

^b Department of Neurosurgery, Allegheny General Hospital, Pittsburgh, PA 15212, United States

^c Department of Statistics, University of Illinois Urbana-Champaign, Urbana, IL 61801, United States

^d Department of Neurology, Carle Foundation Hospital, Urbana, IL 61801, United States

^e Grainger College of Engineering, University of Illinois Urbana-Champaign, Urbana, IL 61801, United States

^f Department of Neurosurgery, Carle Foundation Hospital, Urbana, IL 61801, United States

ARTICLE INFO

Keywords:

Facet arthroplasty
Lumbar spine
Spinal stenosis
Spondylolisthesis
Older adults
Lumbar fusion

ABSTRACT

Background: Facet arthroplasty, an alternative to lumbar fusion, offers stabilization and preserves range of motion. This subanalysis of the TOPS IDE trial (FDA #G160168) compared facet arthroplasty, using the TOPS device, with a standard single-level transforaminal lumbar interbody fusion (TLIF) in patients stratified by age (<65 and ≥65 years) with symptomatic grade 1 degenerative spondylolisthesis with moderate to severe spinal stenosis at L2-5. **Methods:** Patient-reported outcomes (PROMS), including Oswestry disability index (ODI), visual analog pain scales (VAS), and Zurich claudication questionnaires (ZCQ), were assessed at baseline and multiple postoperative timepoints. Radiographic evaluation of flexion/extension range of motion (ROM) occurred at baseline, 12 months, and 24 months. Data were analyzed following an intention-to-treat model. Significance was defined as $p < .05$. **Results:** About 299 patients were included (TOPS=206, TLIF=93). The groups were similar at baseline. At 2 years, the TOPS group had a greater proportion of patients report ≥15-point improvement for ODI (93.8% versus 77.1%, $p = .011$) and ≥20-point improvement for VAS back (84.4% versus 61.8%, $p = .014$). At 1 year, TOPS group had a greater proportion of patients report clinically significant improvements in all ZCQ categories (91.6% versus 78.5%, $p = .012$). In patients <65 years, the TOPS group had improved PROMS compared to TLIF at 2 years; however, these differences were less pronounced in patients ≥65 years old. The TOPS groups preserved more ROM at 12 (2.8° 95%CI [1.87; 3.74], $p < .0001$) and 24 (2.99° 95%CI [1.82; 4.15], $p < .0001$) months compared to TLIF. ROM was similarly preserved in patients aged <65 and ≥65. The rate of adverse events did not differ significantly between treatment groups. **Conclusions:** Facet arthroplasty preserves more ROM in all ages and leads to improved PROMS compared to TLIF, particularly in younger patients.

Background

Lumbar spinal stenosis and spondylolisthesis are a common cause of back pain and indication for surgery [1]. Lumbar fusion is increasingly performed in the United States for the treatment of degenerative disc disease, spondylolisthesis, and stenosis, particularly in the elderly [2–5]. Decompressive surgery is commonly used for both spinal steno-

sis and spondylolisthesis; however, the addition of fusion is debated [6–8]. Fusion is thought to enhance stability and reduce the need for future spinal surgery. In addition to surgeon specific factors, Schneider et al. found higher grade spondylolisthesis and younger patient age to be associated with fusion [9]. Fusion results in decreased range of motion, and this disruption can lead to adjacent segment degeneration [10]. Decompression alone can preserve the range of motion but may

FDA device/drug status: Not applicable.

Author disclosures: **AS:** Nothing to disclose. **AKY:** Nothing to disclose. **AY:** Nothing to disclose. **GH:** Nothing to disclose. **RI:** Nothing to disclose. **PMA:** Nothing to disclose.

* Corresponding author. Carle Neuroscience Institute, 610 N Lincoln Ave, Urbana, IL 61801, USA, Tel.: 1-217-383-6555; fax: 1-217-326-0277.

E-mail address: paul.arnold@carle.com (P.M. Arnold).

<https://doi.org/10.1016/j.xnsj.2024.100329>

Received 24 January 2024; Received in revised form 11 May 2024; Accepted 14 May 2024

Available online 16 May 2024

2666-5484/© 2024 The Authors. Published by Elsevier Inc. on behalf of North American Spine Society. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Table 1
Description and results of prior studies investigating the TOPS device.

Summary of prior studies investigating the TOPS device						
Study	Country	Study type	Patient population	Treatment(s)	PROMS	Radiographic outcomes
Anekstein et al 2015* [15] and Smorgick et al 2019* [12]	Israel	Prospective	Indication: One-level symptomatic lumbar stenosis with grade 1 degenerative spondylolisthesis Level: L4-5 Age: 52-69 years	TOPS (n=10)	Significant decreases in ODI, VAS back, VAS leg at 7 and 11 years; significant increase in SF-36 at 7 and 11 years	Flexion/extension ROM averaged 6.1, 5.08, 4.78, and 4.5 at baseline, 3 months, 1 year, 7 years, and 11 years
Coric et al 2023 [13]	US, Israel	Prospective RCT; TOPS IDE trial [†]	Indication: One-level symptomatic lumbar stenosis with grade 1 degenerative spondylolisthesis Level: L4-5 (95%) Age: 38-80 years	TOPS (n=170), TLIF (n=79)	More TOPS patients achieved MCID for ODI and VAS back at 2 years; insignificant difference in achieving MCID for ZCQ satisfaction and VAS leg at 2 years	TOPS group maintained significantly greater flexion/extension and lateral bending ROM at follow-up
Haleem et al 2021 [16]	United Kingdom	Prospective	Indication: One-level symptomatic lumbar stenosis with grade 1 degenerative spondylolisthesis Level: L3-3 (n=3), L4-5 (n=7) Age: 51-71	TOPS (n=10)	Significant decreases in ODI, VAS back, VAS leg, ZCQ scores at 2 years; significant increase in SF-36 at 2 years	Patients had continued mobility of the stabilized segment at 2 years
Lack et al 2022 [17]	Austria	Prospective	Indication: One-level symptomatic lumbar stenosis with grade 1 degenerative spondylolisthesis Level: L2-3 (n=1), L3-4 (n=6), L4-5 (n=10) Age: 54-82	TOPS (n=17)	Significant decrease in mean VAS back from 8.3 to 1.6 at last follow-up	No significant change in mobility or segmental lordosis at last follow-up
McAfee et al 2007 [14]	Belgium, Brazil, Israel, Turkey	Prospective	Indication: Moderate to severe lumbar stenosis with or without spondylolisthesis Level: L3-4 (n=1), L4-5 (n=28) Age: 52-72	TOPS (n=29)	Mean VAS leg decreased from 88 to 12 at 1 year; mean ODI decreased from 57 to 16 at 1 year; mean ZCQ decreased from 57 to 26 at 1 year	Global motion was preserved in all patients
Pinter et al 2023 [11]	US, Israel	Prospective RCT; TOPS IDE trial [†]	Indication: One-level symptomatic lumbar stenosis with grade 1 degenerative spondylolisthesis Level: L4-5 (95%) Age: 38-79 years	TOPS (n=153)	93.5% achieved MCID for ODI at 1 year; 83.5% and 94.2% achieved MCID for VAS back and leg, respectively, at 1 year; >90% achieved MCID for ZCQ scores at 1 year	Insignificant change in flexion/extension and lateral bending ROM at and above index level at 1 year; significant increase in ROM below index level at 1 year

* Smorgick et al and Anekstein et al published the same patient group at 7 and 11 years follow up.

† The TOPS IDE trial has been published at multiple different timepoints throughout the study. Each publication uses data from the same patient population.

be associated with greater rates of reoperation [7]. Lumbar facet arthroplasty is a proposed method to treat grade I spondylolisthesis with stenosis and aims to restore segmental stability while preserving motion at the index level and adjacent vertebrae [11].

The TOPS system (Premia Spine Ltd., CT, USA) is a total facet replacement device intended for the treatment of degenerative lumbar spondylolisthesis and lumbar stenosis. The system provides stability, the main benefit of fusion, but is motion-preserving, a benefit of decompressive surgery. Application of the device requires a total laminectomy, providing wide decompression, and bilateral total facetectomy. Long-term follow up studies of the implant show sustained declines in leg and back pain and disability, sustained increases in quality of life, and motion preservation (Table 1) [11–17]. The Anatomic Facet Replacement System (AFRS, Facet Solutions Inc., Logan, UT) and the Total Facet Arthroplasty System (TFAS, Archus Orthopedics Inc., Redmond, WA) are similar products. The AFRS was previously shown to decrease pain and claudication scores compared to fusion; however, there were concerns about the device's metal-on-metal design [18]. The TFAS system has very limited clinical data as the IDE trial was discontinued for funding reasons [18].

With the aging population undergoing spinal surgery at increasing rates [3,4], it is critical to evaluate the safety and efficacy of surgical innovations in this population. Prior publications of the TOPS investigational device exemption (IDE) trial have included a limited dataset and did not analyze outcomes by age. Here, we compare the utility of the TOPS system versus standard single level transforaminal lumbar interbody fusion (TLIF) in patients <65 years old (young) and ≥65 years old (old) in 299 patients.

Methods

Study design

This study is part of the TOPS prospective, randomized IDE trial undertaken to evaluate the safety and efficacy of the TOPS device (FDA #G160168). Institutional review board approval was obtained at all participating institutions, and all patients provided written informed consent. Patients aged 35 to 80 years undergoing surgery for symptomatic grade 1 degenerative spondylolisthesis with moderate to severe spinal stenosis and thickening of the ligamentum flavum or scarring of the facet joint capsule at a single level between L2-5 were eligible for inclusion. Patients were randomized to receive the TOPS implant or standard open single-level TLIF. Implantation of the TOPS device has been previously described (Fig. 1) [11]. Surgeons could not be blinded to the treatment group. The patient was initially blinded to the treatment but likely learned their treatment assignment in the follow-up period. Adverse events were recorded by investigators and reviewed by an independent committee who classified their severity (mild, moderate, and severe) and relation to the procedure and device (definitely, probably, possibly, and not related). Detailed information regarding the study methods has been previously described [13].

Data collection

Preoperative demographics (age, height, weight, body mass index (BMI), race, ethnicity), Oswestry Disability Index (ODI), Zurich Claudication Questionnaire (ZCQ), and Visual Analog Scale (VAS) for low

Flexion-Extension X-rays of the TOPS Device

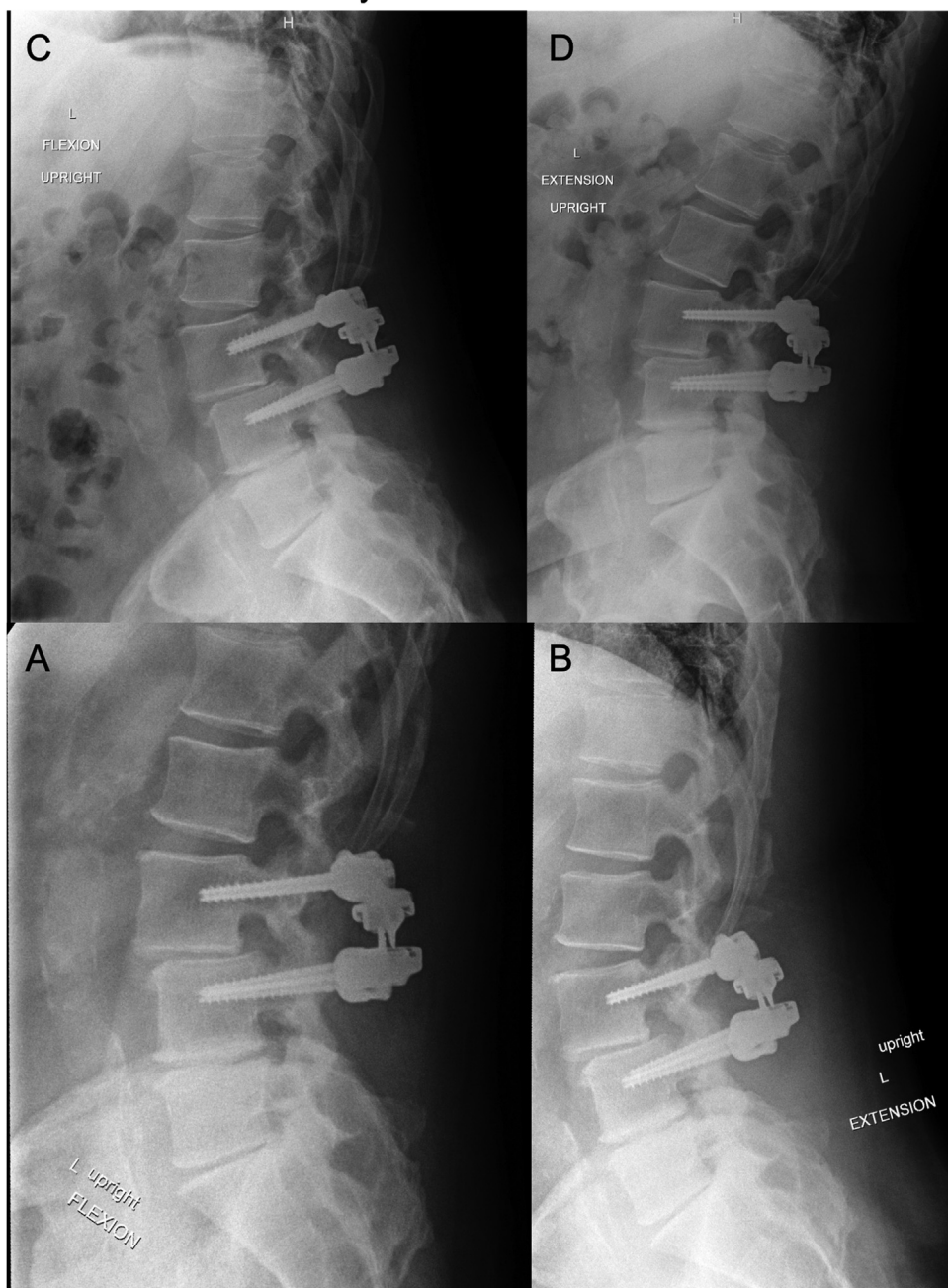


Fig. 1. Flexion-extension lumbar X-rays of the TOPS implant device in a 63-year-old patient at the time of operation (A, B) and 4 years status post operation (C, D). Figure created with BioRender.com.

back and both legs were collected. At postoperative week 6, 3 months, 6 months, 12 months, and 24 months, ODI, ZCQ, and VAS for low back and both legs were repeated. Radiographic measures, flexion-extension range of motion translation (mm) (ROM), flexion-extension ($^{\circ}$ degrees) ROM, and lateral bending ($^{\circ}$ degrees) ROM, were measured preoperatively, at 12 months, and at 24 months. Clinical success was determined at 24 months, and success was defined as ODI reduction ≥ 15 -points in addition to the absence of supplementary surgical intervention, major device adverse events, and new or worsening neurological deficits. Fusion was defined as the presence of bridging trabecular bone across the involved motion segment with angular motion $< 3^{\circ}$ and translational motion < 2 mm. If bony bridging was indeterminate/unable to assess, fusion was assumed to have not occurred. Patients are being followed for up to 48 months, however, the current sample sizes are too small at later timepoints for productive analysis.

Statistical analysis

Demographic data was summarized using standard summary statistics and is presented as mean (standard deviation) or number of patients (N, percent of cohort). Differences in continuous variables were calculated with Student's T-Test. Cohen's *d* is reported for effect size of continuous variables. Fisher's exact test was used for differences in categorical variables. Differences in adverse were calculated with z-test for proportions and reported as mean difference in proportion (%) with a 95% confidence interval. p-values $< .05$ are considered significant. Minimal clinically important differences (MCID) for each variable are ≥ 15 -point improvement ODI [13], ≥ 0.5 -point improvement ZCQ Symptoms and Physical [19], ≤ 2.5 total ZCQ Satisfaction, and ≥ 20 -point improvement VAS worst leg and low back [20,21]. Patients were also stratified by age with < 65 years considered young and ≥ 65

years considered old. Data was analyzed following an intention-to-treat model.

Results

Two hundred ninety-nine patients were included, with 206 patients receiving the TOPS implant (68.9%) and 93 undergoing TLIF (31.1%). The mean age was 63.3 (8.2) years and 63.9 (8.6) years in the TOPS and TLIF groups, respectively. In patients ≥ 65 years old, the mean age was 70.6 (3.8) years and 70.6 (4.0) years in the TOPS and TLIF groups, respectively. The mean BMI in all groups was in the overweight or obese ranges. Most patients were white, very few had prior lumbar surgeries, and the majority were non-smokers. Nearly all patients underwent surgery at L4-5 (95.1% TOPS, 93.5% TLIF) with remaining patients undergoing the operation at L3-4. No patients underwent surgery at L2-3.

At baseline, there were no significant differences between the TOPS and TLIF groups in ODI, ZCQ-Symptom Score, ZCQ-Physical Score, VAS Worst Leg, or VAS Back for any age grouping. Baseline scores are summarized by age group in Table 2. All ROM measurements were similar at the index level, below index level, and above index level for all groups, except in the older group where flexion-extension translation at the index level was greater in the TLIF group (0.93 mm versus 1.43 mm, $p=.009$).

Patient-reported outcome measures

Oswestry Disability Index (ODI)

At week 6 and 3 months the ODI was significantly lower for the entire TOPS group ($p=.002$ [$d=0.41$], 0.005 [$d=0.38$], respectively) and the old TOPS group ($p=.008$ [$d=0.49$], 0.049 [$d=0.36$], respectively) compared to TLIF (Table 2). The entire TOPS group again had significantly lower ODI scores at months 12 and 24 ($p=.019$ [$d=0.35$], 0.001 [$d=0.69$], respectively), but no significant differences in the old groups were observed. At 2 years, the entire TOPS and the young TOPS group also had significantly more patients report at least a 15-point improvement on ODI (93.8% vs 77.1%, $p=.001$; 94.1% vs 64.7%, $p=.006$). The proportion of patients achieving MCID for ODI did not differ at any timepoint when patients were compared by age within their respective treatment group ($p>.05$).

Visual Analog Scale (VAS)

For worst leg VAS, the young TOPS group reported significantly less pain at 12 and 24 months (12.4 versus 23.6, $p=.031$, $d=0.46$; 8.7 versus 28.3, $p=.006$, $d=0.81$, respectively) (Table 2). Over 80% of all patients reported an improvement of ≥ 20 -points at all follow-up times with no significant differences between TOPS and TLIF. For low back VAS, the TOPS group had significantly lower pain scores at all postoperative timepoints ($p<.05$, $d>0.31$). Additionally, the TOPS group had more patients report improvements of ≥ 20 -points at week six and 24 months (83.5% vs 68.7%, $p=.009$; 84.4% vs 61.8%, $p=.014$, respectively). This difference was also significant in the young but not the old groups. The proportion of patients achieving MCID for worst leg or low back VAS did not differ at any timepoint when patients were compared by age within their respective treatment group ($p>.05$).

Zurich Claudication Questionnaire (ZCQ)

At week six, 12 months, and 24 months, the ZCQ symptom score was significantly lower in the TOPS group compared to TLIF ($p=.038$ [$d=0.27$], 0.038 [$d=0.31$], 0.001 [$d=0.52$], respectively) (Table 2). At 24 months a greater proportion the younger TOPS group reported ZCQ symptom score improvement ≥ 0.5 -points (96.1% vs 76.5%, $p=.031$), but no other differences in the proportion of patients reporting MCID were observed between the groups. The proportion of patients reporting MCID in ZCQ physical score was significantly greater at 6 weeks for the TOPS group (86.7% vs 76.2%, $p=.036$) and at 12 months for the overall (96.5% vs 84.6%, $p=.007$), young (96.2% vs 83.9%, $p=.04$), and old (96.9% vs 85.3%, $p=.045$) TOPS groups compared to TLIF. At week 6,

12 months, and 24 months the ZCQ satisfaction score was significantly decreased (improved) in the overall and young TOPS groups.

The proportion of patients reporting MCID in all categories of ZCQ was significantly greater at 1 year in the overall TOPS group (91.6% vs 78.5%, $p=.012$) and the young TOPS group (92.3% vs 74.2%, $p=.022$). This proportion was not significantly different at any timepoint when patients were compared by age within their respective treatment groups.

Radiographic parameters

Flexion-extension range of motion (ROM)

At the index level the preoperative mean flexion-extension ROM was similar for both treatment groups (4.14° (3.06) and 4.47° (3.42) for TOPS and TLIF ($p=.4096$), respectively) (Fig. 2). From preoperative to 12 and 24 months there were no significant differences in the index level ROM for the TOPS groups, however, there was a significant decline in index level ROM for the TLIF groups ($p<.0001$). The TOPS groups preserved significantly more ROM at 12 (estimated mean difference (EMD) 2.8° 95%CI[1.87; 3.74], $p<.0001$, $d=0.86$) and 24 (EMD 2.99° 95%CI[1.82; 4.15], $p<.0001$, $d=0.93$) months postoperatively versus TLIF. ROM was similarly preserved in the young (12 months: EMD 2.65° 95%CI[1.41; 3.90], $p=.0001$, $d=0.88$; 24 months: EMD 3.30° 95%CI[1.61; 4.99], $p=.002$, $d=1.03$) and old (12 months: EMD 2.96° 95%CI[1.53; 4.39], $p=.0001$, $d=0.84$; 24 months: EMD 2.64° 95%CI[0.98; 4.31], $p=.0023$, $d=0.81$). Comparing the young and old patients within the TOPS group, there was no significant difference in index level ROM at any timepoint.

Below the index level the mean preoperative ROM for TOPS and TLIF groups were 5.39° (3.63) and 5.70° (3.69) ($p=.5073$), respectively. There was a significant gain in ROM for the TOPS group at 12 and 24 months ($p<.0001$). Similarly, the young and old TOPS groups significantly gained ROM below the index level. The TLIF groups insignificantly gained ROM at 12 and 24 months. At 12 months postoperative the TOPS group gained significantly more ROM compared to TLIF (EMD 1.24° 95%CI[0.05; 2.43], $p=.0413$, $d=0.30$). Within the TOPS group, the younger patients gained significantly more ROM below the index level at 12 (EMD 1.86° 95%CI[0.56; 3.17], $p=.006$, $d=0.46$) and 24 months (EMD 2.31° 95%CI[0.79; 3.83], $p=.004$, $d=0.57$).

Above the index level the mean preoperative ROM for TOPS and TLIF groups were 3.45° (3.06) and 3.58° (3.06) ($p=.7376$), respectively. There was a significant gain in ROM for the overall (12 months, $p=.0064$; 24 months, $p=.0101$) and young TOPS groups (12 months, $p=.0437$; 24 months, $p=.0155$), however, there was no significant difference for the old TOPS group. The TLIF groups insignificantly gained ROM at both timepoints. At both timepoints, there were no significant differences in ROM between the TOPS and TLIF groups. Comparing the young and old patients within the TOPS group, there was no significant difference in ROM above the index level at any timepoint.

Lateral bending angular motion (AM) at the index level

At baseline there were no significant differences in AM between TOPS and TLIF groups. The TOPS group preserved significantly more AM compared to TLIF at 12 (EMD 2.53° 95%CI[1.70; 3.37], $p<.0001$, $d=0.08$) and 24 (EMD 2.51° 95%CI[1.21; 3.82], $p<.002$, $d=0.08$) months. There were no significant changes in AM from baseline to 12 or 24 months in the TOPS groups. However, the TLIF groups lost AM at 12 (-2.15° (1.90), $p<.0001$) and 24 (-2.19° (1.75), $p<.0002$) months. Comparing the young and old patients within the TOPS group, there was no significant difference in AM at the index level at any timepoint.

Surgical outcomes

Operation

Across all groups there were no significant difference in operating time, length of hospital stay, and estimated blood loss (EBL). The mean

Table 2
Patient reported outcome measures (PROM) at baseline, 6 weeks, and at months 3, 6, 12, and 24.

Patient reported outcome measures													
Age group	Time interval	Group	N	ODI		VAS-worst leg pain		VAS-low back pain		ZCQ			
				Mean (SD)	MCID	Mean (SD)	MCID	Mean (SD)	MCID	Symptom, MCID	Physical, MCID	Satisfaction, MCID	
All Patients	Preoperative	TOPS	206	56.5 (12.1)		82.7 (13.5)		68.6 (23.3)					
		Fusion	93	55.8 (13.1)		85.1 (10.8)		69.5 (22.2)					
	6 weeks	TOPS	194	23.5 (16.4)*	84.0	12.9 (20.5)	92.8	18.5 (18)*	83.5*	93.8		86.7*	96.4
		Fusion	84	30.2 (17)	73.8	17.9 (25.2)	92.8	27.7 (25.3)	68.7	90.5		76.2	95.2
	Month 3	TOPS	183	15.7 (16.5)*	89.1	13.3 (22.5)	94.5	16.2 (21.3)*	83.6	95.6		91.2	94
		Fusion	82	22.1 (17.8)	84.1	15.9 (23.7)	92.7	23.1 (24.5)	79.3	96.3		91.4	95.1
	Month 6	TOPS	171	13.4 (15.5)	91.8	12.9 (22.7)	92.4	14.7 (21.1)*	86	95.9		90.6	94.7
		Fusion	74	16.0 (15.9)	90.5	17 (24.9)	91.9	22.7 (24.8)	79.7	93.2		91.9	95.9
	Month 12	TOPS	143	11.6 (13.7)*	94.4	12.8 (22)	94.4	12.4 (19.6)*	86	96.5		96.5*	94.4
		Fusion	65	16.9 (17.2)	89.2	18.7 (27.8)	90.8	24.5 (27.6)	76.9	92.3		84.6	86.2
	Month 24	TOPS	96	9.4 (14.5)*	93.8*	13.7 (24.2)	90.6	11.1 (18.1)*	84.4*	93.8		92.7	92.7
		Fusion	35	21.1 (22.3)	77.1	23.3 (33.8)	88.2	30.9 (33.1)	61.8	85.7		82.9	88.6
	Patients <65 years	Preoperative	TOPS	111	56.8 (12.2)		82.7 (13.3)		72 (23.2)				
			Fusion	43	56.5 (12.4)		85.3 (12.2)		71.1 (22.4)				
6 weeks		TOPS	106	24.7 (16.6)	84.0	13.7 (21.6)	89.6	19.8 (18.2)	84*	94.3		86.9	96.3
		Fusion	38	30.5 (17.4)	78.9	20.2 (27.6)	91.9	27.4 (24.8)	64.9	86.8		76.3	92.1
Month 3		TOPS	100	15.3 (16.6)	89.0	11.9 (21.2)	95	18 (22.8)	82	94.9		93	97
		Fusion	36	21.9 (17.8)	80.6	15.1 (22.9)	91.7	25.8 (25.3)	72.2	100		97.1	97.1
Month 6		TOPS	90	13.7 (16.0)	90.0	12.9 (22.9)	91.1	15.6 (22.6)	86.7	97.8		92.2	95.6
		Fusion	34	16.8 (15.9)	91.2	14.6 (23.1)	91.2	23.3 (23.6)	76.5	97.1		91.2	97.1
Month 12		TOPS	78	11.8 (14.9)	92.3	12.4 (20)*	94.9	13.4 (19.4)*	85.9	97.4		96.2*	93.6
		Fusion	31	18.1 (17.9)	87.1	23.6 (32.6)	83.9	28.6 (30.9)	71	90.3		83.9	80.6
Month 24		TOPS	51	7 (13.1)*	94.1*	8.7 (17)*	96.1	9.5 (18.7)*	86.3*	96.1*		96.1	94.1
		Fusion	17	26.4 (25.6)	64.7	28.3 (39.4)	81.3	35.3 (35.5)	50	76.5		70.6	76.5
Patients ≥ 65 years		Preoperative	TOPS	95	56.2 (12.0)		82.7 (13.8)		64.6 (22.8)				
			Fusion	50	55.2 (13.7)		84.9 (9.6)		68.1 (22.2)				
	6 weeks	TOPS	88	22.0 (16.2)*	84.1	12 (19.2)	96.6	16.9 (17.6)*	83	93.2		86.4	96.6
		Fusion	46	30.0 (16.8)	69.6	16 (23.3)	93.5	27.9 (26)	71.7	93.5		76.1	97.8
	Month 3	TOPS	83	16.1 (16.5)*	89.2	15 (23.9)	93.9	14.2 (19.2)	85.5	96.3		89	90.2
		Fusion	46	22.3 (18.0)	87.0	16.6 (24.5)	93.5	20.9 (23.8)	84.8	93.5		87	93.5
	Month 6	TOPS	81	13.1 (15.0)	93.8	13 (22.6)	93.8	13.8 (19.3)*	85.2	93.8		88.9	93.8
		Fusion	40	17.0 (16.2)	90.0	19.1 (26.4)	92.5	22.2 (26)	82.5	90		92.5	95
	Month 12	TOPS	65	11.3 (12.3)	96.9	13.2 (24.3)	93.8	11.2 (19.9)*	86.2	95.4		96.9*	95.4
		Fusion	34	15.8 (16.7)	91.2	14.2 (22.1)	97.1	20.8 (24)	82.4	94.1		85.3	91.2
	Month 24	TOPS	45	12.1 (15.6)	93.3	19.4 (29.5)	84.4	13 (17.3)	82.2	91.1		88.9	91.1
		Fusion	18	16.0 (17.8)	88.9	18.8 (28.5)	94.4	27 (31.3)	72.2	94.4		94.4	100

Abbreviations: MCID, minimal clinically important difference; N, number of patients; ODI, Oswestry Disability Index; SD standard deviation; TOPS, total posterior spine system; VAS, Visual Analog Scale; ZCQ, Zurich Claudication Questionnaire.

Minimal clinically important difference (MCID) is reported as a percentage of the cohort. MCID for each PROM are ≥15-point improvement ODI, ≥0.5-point improvement ZCQ Symptoms and Physical, ≤2.5 total ZCQ Satisfaction, and ≥20-point improvement VAS worst leg and low back.

* Significant difference (p<.05) between TOPS and fusion groups by unpaired T-test (means) or Fisher's exact test (MCID).

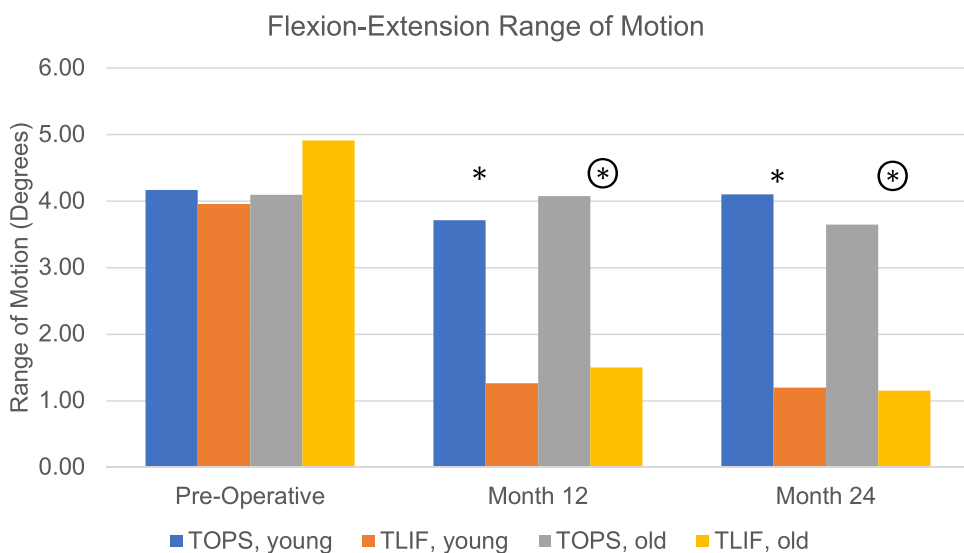


Fig. 2. Flexion-extension range of motion in degrees for the index level at baseline and at 12 and 24 months status post operation where * denotes a significant difference between the young TOPS and TLIF groups and ⊗ denotes a significant difference between the old TOPS and TLIF groups ($p < .05$). Color coding: TOPS <65 years (blue), TLIF <65 years (orange), TOPS ≥65 years (grey), TLIF ≥65 years (yellow).

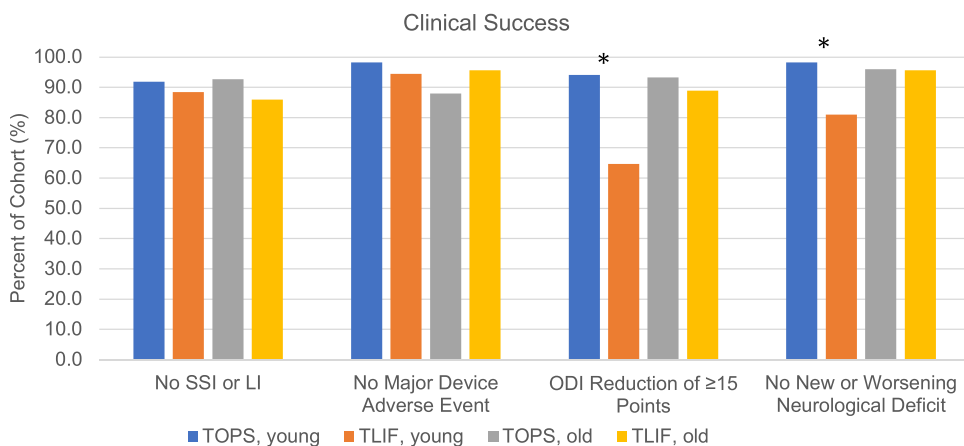


Fig. 3. Percent of cohort meeting clinical success criteria, where * denotes a significant difference between the young TOPS and TLIF groups. Color coding: TOPS <65 years (blue), TLIF <65 years (orange), TOPS ≥65 years (grey), TLIF ≥65 years (yellow).

operating time for the TOPS and TLIF groups was 181.7 (57.0) minutes and 176.9 (56.7) minutes, respectively ($p = .495$). Mean length of stay for the TOPS and TLIF groups was 2.86 (3.62) days and 2.86 (1.75) days, respectively ($p = .997$). The mean EBL for the TOPS and TLIF groups was 199.6 (146) mL and 214.8 (133.4) mL, respectively ($p = .395$). EBL was greater in the older TLIF versus younger TLIF group (173.8 (122.0) mL versus 249.9 (133.9) mL, $p = .005$), but no significant difference was seen in the TOPS groups ($p = .059$). Operating time and length of stay were insignificantly greater in older TOPS and TLIF patients.

Clinical Success

Major device-related adverse events at two years postoperative, including breakage, disassembly, screw loosening, or increase in spondylolisthesis grade, were seen in 6.67% and 4.88% of TOPS and TLIF patients ($p = .686$), respectively (Fig. 3). New or worsening neurological deficits were observed in 2.80% and 11.36% of TOPS and TLIF patients at two years postoperative (-8.56% 95%CI[-18.45; 1.33], $p = .033$), respectively (Fig. 3). In patients <65 years, significantly more deficits were observed in the TLIF patients (-17.29% 95%CI[-34.43; -0.16], $p = .006$). Fusion status failure occurred in 43.90% of TLIF patients at 2 years postoperative. Supplemental surgical interventions, including reoperation and lumbar injections, were performed in 7.77% and 12.90% of TOPS and TLIF patients at 2 years postoperative ($p = .158$), respectively (Fig. 3). The use of supplemental interventions was similar between young and old TOPS patients ($p = .843$).

Safety profile

Adverse events occurred in 65.2% of TOPS patients and 61.5% of TLIF patients (-3.78% 95%CI[-7.89; 15.45], $p = .609$), with 25.2% and 16.7% being serious (8.57% 95%CI[-0.92; 18.06], $p = .129$), respectively. Device-related adverse events occurred in 21.4% of TOPS patients and 28.1% of TLIF patients (-6.70% 95%CI[-17.26; 3.87], $p = .256$). Device-related serious adverse events occurred in 4.3% and 5.2% of TOPS and TLIF patients, respectively (-0.92% 95%CI[-6.14; 4.30], $p = .949$). Three deaths (1.4%) in the TOPS group occurred, and no deaths were thought to be related to the intervention. No deaths occurred in the TLIF group. There were greater adverse events seen in the old versus young TOPS groups (21.81% 95%CI[9.37; 34.26], $p = .002$). However, the adverse events related to the TOPS device and procedure were insignificantly different in the old versus young (9.15% 95%CI[-2.00; 20.29], $p = .149$ (device); 13.32% 95%CI[-0.10; 26.75], $p = .074$ (procedure), respectively). The frequency of adverse events related to hardware and the procedure were similar for the old and young TLIF groups.

Discussion

The present study provides a comparative analysis of the TOPS IDE study with two treatment arms, TOPS and TLIF. Prior reports from the trial have included a limited dataset [13]. This analysis also compares the clinical outcomes of both treatments in younger versus older pa-

tients. The results suggest the TOPS device performs favorably compared to traditional TLIF in treatment of single-level grade 1 spondylolisthesis and spinal stenosis. PROMS were significantly greater in the TOPS group at multiple timepoints, and ROM was better preserved with TOPS. We find differences in PROMS between treatments are less pronounced in older patients, indicating that while TOPS does improve PROMS significantly from baseline in this population, TOPS may not outperform TLIF in older patients.

Treatment of spondylolisthesis with spinal stenosis is debated, and both operative and non-operative therapies are available. The Spine Patient Outcomes Research Trial (SPORT) found operatively treated patients had significantly greater improvements in pain and function after 8 years compared to nonoperative treatment [22]. Similarly, the Maine Lumbar Spine Study (MLSS) found greater improvement in surgical patients after four years; however, the study also found that after 8 to 10 years the surgical and non-surgical groups had similar pain and function [23]. Spinal fusion is frequently used to treat degenerative spondylolisthesis but decompression-only is also an option [8]. While fusion has been thought to reduce the need for future reoperations, the NORDSTEN-DS trial by Austovell et al did not identify significant differences in reoperation rates nor patient-reported outcomes when compared to decompression-only [24]. Similarly, the Swedish Spinal Stenosis Study also did not identify differences in patient-reported outcomes or reoperation for lumbar stenosis with or without spondylolisthesis [8]. However, a meta-analysis of patients ≥ 65 years with lumbar spinal stenosis and degenerative spondylolisthesis found that patients who underwent decompression only had fewer surgical complications but had worsening degenerative spondylolisthesis postoperatively [25]. Clinical trials comparing lumbar facet arthroplasty with decompression with or without fusion in do not exist for the elderly.

Spinal fusion stabilizes the damaged segment while significantly restricting segmental ROM. One and two-level fusion causes significant decreases in trunk forward flexion, although the effect of this on activities of daily living is not known [26]. While the loss of ROM may not be noticeable in daily living, adjacent segment degeneration caused by the uneven distribution of forces can be seen radiographically [10]. As shown in the present study, usage of the TOPS device preserves greater segmental ROM and can improve adjacent segment ROM compared to fusion [11–17]. The wider decompression and motion preservation likely contributes to improved patient reported outcomes in this study and others [11–17]. Similar to the present study, Coric et al reported more TOPS patients with ≥ 15 -point ODI improvement and ≥ 20 -point improvement in VAS back pain compared to the fusion group at 24 months [13]. In Coric et al and the present study, the proportion of TOPS patients meeting MCIDs exceeded both the fusion and decompression-only groups of the NORDSTEN-DS trial [13,24].

Spinal surgery for lumbar stenosis and spondylolisthesis is increasingly performed, particularly in the elderly population. As age increases, patients undergoing spinal surgery are more likely to have greater comorbidities and a more fragile physical condition [27,28]. A meta-analysis of surgical outcomes for lumbar stenosis and spondylolisthesis comparing young (< 80) versus old (≥ 80) found significantly greater postoperative complications, including mortality and reoperation, occurred in the elderly for both decompression-only and decompression with fusion [27]. However, the association of age and complications is not fully elucidated with some studies only finding differences in minor complications [28,29] and others finding no significant differences [30]. Additionally, it is worth noting that older patients may require more extensive surgery [30], and the presence of comorbidities is a risk factor for complications and mortality independent of age [28,31]. Despite the potentially greater risk of surgery, patient reported outcomes in older adults do not significantly differ from their younger counterparts in showing a difference from baseline or compared to a non-surgical treatment group [27–30,32].

While the TOPS IDE trial and other facet arthroplasty device studies have excluded patients aged > 80 years, it is important to evaluate facet

arthroplasty's safety and functionality in older adults. The age range in the present study was 38 to 80, with a median of 70 years in the older group. Similar to previous studies, we found significantly more adverse events in the older TOPS group, however, there was not a significant difference in device-related or procedure-related adverse events by age in either treatment group. The greater occurrence of adverse events in the elderly may be more attributable to comorbidities rather than the device or procedure.

Similar to prior spinal surgery studies we find both the young and old patients have significant improvements from baseline with both facet arthroplasty and TLIF. For many PROMS facet arthroplasty outperforms TLIF. However, fewer significant differences in the PROMS of TOPS versus TLIF patients are observed when only the older population is considered. ROM at the index level was preserved for all ages with TOPS, however, the gain of adjacent level ROM was not observed in older patients. Prior facet arthroplasty studies have not stratified patients by age making it challenging to determine expected treatment responses in the older population.

The primary limitation of this study is the limited long-term follow-up. At six and 12 months, approximately 80% and 69%, respectively, of all patients reported outcomes while the remaining patients have yet to reach these timepoints. Less than half of the cohort was available for follow-up at 24 months (48.6% TOPS, 37.6% TLIF), and the cohort further declined to less than a quarter of the cohort at 36 and 48 months (23.79%, 6.31% TOPS; 14%, 5.4% TLIF). Thus, comparisons between treatments after 12 months should be considered in the context of a limited sample, and this limited sample introduces potential bias in results. The percentage of patients reporting long-term study outcomes is less than previously seen in the SPORT [22] and MLSS [23] trials due to the interim nature of this analysis; continued follow-up with adequate retention is necessary to better understand the impact of the TOPS device. The sample size also limits the understanding of long-term complications, including adjacent segment disease. Smorgick et al published follow up data for 11 years and did not identify significant complications in later years; however, the sample size was limited to only 10 patients [12]. Additionally, this study excluded patients over 80 years old, which may limit its generalizability to this population. Lastly, the rate of pseudarthrosis, or fusion failure, was relatively high in the TLIF group compared to the 5%–35% pseudarthrosis rate observed in prior studies [33]. The exact cause is unknown, however, variation in pseudarthrosis rates across studies may be attributed to lack of a universally accepted quantitative definition of fusion, differences in timing of imaging and imaging modality, and poor inter-rater reliability [33,34].

Conclusions

This subanalysis of the TOPS IDE trial compared facet arthroplasty with TLIF for the treatment of single-level symptomatic grade 1 degenerative spondylolisthesis with spinal stenosis in younger (< 65 years) and older adults (≥ 65). TOPS performed equally or better than TLIF for PROMS at multiple timepoints, which may be due to the wider decompression and motion preservation. The benefit of TOPS over TLIF was more apparent in younger patients, however, patients of all ages achieved significant improvement in PROMS from baseline.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

None.

References

- [1] Katz JN, Zimmerman ZE, Mass H, Makhni MC. Diagnosis and management of lumbar spinal stenosis: a review. *JAMA* 2022;327:1688–99. doi:10.1001/jama.2022.5921.
- [2] Yoshihara H, Yoneoka D. National trends in the surgical treatment for lumbar degenerative disc disease: United States, 2000 to 2009. *Spine J* 2015;15:265–71. doi:10.1016/j.spinee.2014.09.026.
- [3] O'Lynnner TM, Zuckerman SL, Morone PJ, et al. Trends for spine surgery for the elderly: implications for access to healthcare in North America. *Neurosurgery* 2015;77(Suppl 4):S136–41. doi:10.1227/NEU.0000000000000945.
- [4] Deyo RA, Gray DT, Kreuter W, Mirza S, BI Martin. United States trends in lumbar fusion surgery for degenerative conditions. *Spine (Phila Pa 1976)* 2005;30:1441–5 discussion 1446–7. doi:10.1097/01.brs.0000166503.37969.8a.
- [5] Saleh A, Thirukumaran C, Mesfin A, Molinari RW. Complications and readmission after lumbar spine surgery in elderly patients: an analysis of 2,320 patients. *Spine J* 2017;17:1106–12. doi:10.1016/j.spinee.2017.03.019.
- [6] Matz PG, Meagher RJ, Lamer T, et al. Guideline summary review: an evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spondylolisthesis. *Spine J* 2016;16:439–48. doi:10.1016/j.spinee.2015.11.055.
- [7] Ghogawala Z, Dziura J, Butler WE, et al. Laminectomy plus fusion versus laminectomy alone for lumbar spondylolisthesis. *N Engl J Med* 2016;374:1424–34. doi:10.1056/NEJMoa1508788.
- [8] Forsth P, Olafsson G, Carlsson T, et al. A randomized, controlled trial of fusion surgery for lumbar spinal stenosis. *N Engl J Med* 2016;374:1413–23. doi:10.1056/NEJMoa1513721.
- [9] Schneider N, Fisher C, Glennie A, et al. Lumbar degenerative spondylolisthesis: factors associated with the decision to fuse. *Spine J* 2021;21:821–8. doi:10.1016/j.spinee.2020.11.010.
- [10] Porrino J, Rao A, Moran J, et al. Current concepts of spondylolysis and posterior spinal motion preservation for radiologists. *Skeletal Radiol* 2021;50:2169–84. doi:10.1007/s00256-021-03840-6.
- [11] Pinter ZW, Freedman BA, Nassr A, et al. A prospective study of lumbar facet arthroplasty in the treatment of degenerative spondylolisthesis and stenosis: results from the total posterior spine system (TOPS) IDE study. *Clin Spine Surg* 2023;36:E59–69. doi:10.1097/BSD.0000000000001365.
- [12] Smorgick Y, Mirovsky Y, Floman Y, et al. Long-term results for total lumbar facet joint replacement in the management of lumbar degenerative spondylolisthesis. *J Neurosurg Spine* 2019;32(1):1–6. doi:10.3171/2019.7.SPINE19150.
- [13] Coric D, Nassr A, Kim PK, et al. Prospective, randomized controlled multicenter study of posterior lumbar facet arthroplasty for the treatment of spondylolisthesis. *J Neurosurg Spine* 2023;38:115–25. doi:10.3171/2022.7.SPINE22536.
- [14] McAfee P, Khoo LT, Pimenta L, et al. Treatment of lumbar spinal stenosis with a total posterior arthroplasty prosthesis: implant description, surgical technique, and a prospective report on 29 patients. *Neurosurg Focus* 2007;22:E13. doi:10.3171/foc.2007.22.1.13.
- [15] Anekstein Y, Floman Y, Smorgick Y, et al. Seven years follow-up for total lumbar facet joint replacement (TOPS) in the management of lumbar spinal stenosis and degenerative spondylolisthesis. *Eur Spine J* 2015;24:2306–14. doi:10.1007/s00586-015-3850-0.
- [16] Haleem S, Ahmed A, Ganesan S, McGillion SF, Fowler JL. Mean 5-year follow-up results of a facet replacement device in the treatment of lumbar spinal stenosis and degenerative spondylolisthesis. *World Neurosurg* 2021;152:e645–51. doi:10.1016/j.wneu.2021.06.045.
- [17] Lack W, Kutschera HP, Krugluger J. Mobility-maintaining facet arthroplasty of the lumbar spine with the second-generation TOPS system: a case series. *Oper Neurosurg (Hagerstown)* 2022;23:14–21. doi:10.1227/ons.0000000000000226.
- [18] Gu BJ, Blue R, Yoon J, Welch WC. Posterior lumbar facet replacement and arthroplasty. *Neurosurg Clin N Am* 2021;32:521–6. doi:10.1016/j.nec.2021.05.011.
- [19] Fukushima M, Oka H, Oshima Y, et al. Evaluation of the minimum clinically important differences of the Zurich Claudication Questionnaire in patients with lumbar spinal stenosis. *Clin Spine Surg* 2020;33:E499–503. doi:10.1097/BSD.0000000000000983.
- [20] Hagg O, Fritzell P, Nordwall A, Swedish Lumbar Spine Study G. The clinical importance of changes in outcome scores after treatment for chronic low back pain. *Eur Spine J* 2003;12:12–20. doi:10.1007/s00586-002-0464-0.
- [21] Kovacs FM, Abaira V, Royuela A, et al. Minimal clinically important change for pain intensity and disability in patients with nonspecific low back pain. *Spine (Phila Pa 1976)* 2007;32:2915–20. doi:10.1097/BRS.0b013e31815b75ae.
- [22] Abdu WA, Sacks OA, Tosteson ANA, et al. Long-term results of surgery compared with nonoperative treatment for lumbar degenerative spondylolisthesis in the spine patient outcomes research trial (SPORT). *Spine (Phila Pa 1976)* 2018;43:1619–30. doi:10.1097/BRS.0000000000002682.
- [23] Atlas SJ, Keller RB, Wu YA, Deyo RA, Singer DE. Long-term outcomes of surgical and nonsurgical management of lumbar spinal stenosis: 8 to 10 year results from the maine lumbar spine study. *Spine (Phila Pa 1976)* 2005;30:936–43. doi:10.1097/01.brs.0000158953.57966.c0.
- [24] Austevoll IM, Hermansen E, Fagerland MW, et al. Decompression with or without fusion in degenerative lumbar spondylolisthesis. *N Engl J Med* 2021;385:526–38. doi:10.1056/NEJMoa2100990.
- [25] Abdel-Fattah AR, Bell F, Boden L, et al. To fuse or not to fuse: the elderly patient with lumbar stenosis and low-grade spondylolisthesis. Systematic review and meta-analysis of randomised controlled trials. *Surgeon* 2023;21:e23–31. doi:10.1016/j.surge.2022.02.008.
- [26] Stief F, Meurer A, Wienand J, Rauschmann M, Rickert M. Has a mono- or bisegmental lumbar spinal fusion surgery an influence on self-assessed quality of life, trunk range of motion, and gait performance? *Spine (Phila Pa 1976)* 2015;40:E618–26. doi:10.1097/BRS.0000000000000885.
- [27] Liang H, Lu S, Jiang D, Fei Q. Clinical outcomes of lumbar spinal surgery in patients 80 years or older with lumbar stenosis or spondylolisthesis: a systematic review and meta-analysis. *Eur Spine J* 2020;29:2129–42. doi:10.1007/s00586-019-06261-1.
- [28] Rault F, Briant AR, Kamga H, Gaberel T, Emery E. Surgical management of lumbar spinal stenosis in patients over 80: is there an increased risk? *Neurosurg Rev* 2022;45:2385–99. doi:10.1007/s10143-022-01756-w.
- [29] Giannadakis C, Solheim O, Jakola AS, et al. Surgery for lumbar spinal stenosis in individuals aged 80 and older: a multicenter observational study. *J Am Geriatr Soc* 2016;64:2011–18. doi:10.1111/jgs.14311.
- [30] Rihn JA, Hilibrand AS, Zhao W, et al. Effectiveness of surgery for lumbar stenosis and degenerative spondylolisthesis in the octogenarian population: analysis of the Spine Patient Outcomes Research Trial (SPORT) data. *J Bone Joint Surg Am* 2015;97:177–85. doi:10.2106/JBJS.N.00313.
- [31] Li G, Patil CG, Lad SP, et al. Effects of age and comorbidities on complication rates and adverse outcomes after lumbar laminectomy in elderly patients. *Spine (Phila Pa 1976)* 2008;33:1250–5. doi:10.1097/BRS.0b013e3181714a44.
- [32] Shabat S, Arinon Z, Folman Y, et al. Long-term outcome of decompressive surgery for Lumbar spinal stenosis in octogenarians. *Eur Spine J* 2008;17:193–8. doi:10.1007/s00586-007-0514-8.
- [33] Chun DS, Baker KC, Hsu WK. Lumbar pseudarthrosis: a review of current diagnosis and treatment. *Neurosurg Focus* 2015;39:E10. doi:10.3171/2015.7.FOCUS15292.
- [34] Benson JC, Lehman VT, Sebastian AS, et al. Successful fusion versus pseudarthrosis after spinal instrumentation: a comprehensive imaging review. *Neuroradiology* 2022;64:1719–28. doi:10.1007/s00234-022-02992-z.