ELSEVIER

The Spine Journal ■■ (2017) ■■–■■

THE SPINE JOURNAL

Clinical Study

Identifying subsets of patients with single-level degenerative disc disease for lumbar fusion: the value of prognostic tests in surgical decision making

Victor E. Staartjes, BMed[a,b,*], Pieter-Paul A. Vergroesen, MD, PhD[c,d], Dick J. Zeilstra, MD, PhD[a], Marc L. Schröder, MD, PhD[a]

[a]*Department of Neurosurgery, c/o Bergman Clinics, Rijksweg 69, 1411 GE Naarden, The Netherlands*
[b]*Faculty of Medicine, University of Zurich, Pestalozzistrasse 3, CH 8091 Zurich, Switzerland*
[c]*Department of Orthopedic Surgery, Noordwest Ziekenhuisgroep, Wilhelminalaan 12, 1815 JD Alkmaar, The Netherlands*
[d]*Amsterdam Movement Sciences Institute, VU Amsterdam, Van der Boechorststraat 7, 1081 BT Amsterdam, The Netherlands*

Received 15 March 2017; revised 8 July 2017; accepted 15 August 2017

Abstract

BACKGROUND CONTEXT: Fusion surgery for degenerative disc disease (DDD) has become a standard of care, albeit not without controversy. Outcomes are inconsistent and a superiority over conservative treatment is debatable. Proper patient selection is key to clinical success, and a comprehensive understanding of prognostic tests does not currently exist.

PURPOSE: This study aimed to investigate the value of prognostic tests and sociodemographic factors in predicting outcomes following lumbar fusion surgery for DDD.

STUDY DESIGN: This is a retrospective analysis of prospectively collected data.

PATIENT SAMPLE: We included patients who underwent fusion surgery for DDD between 2010 and 2016.

OUTCOME MEASURES: The outcome measures included pre- and postoperative visual analog scale and Oswestry Disability Index scores.

MATERIALS AND METHODS: Prospectively collected patient data were reviewed for preoperative tests, perioperative data, and clinical outcomes. Prognostic tests used were discography, pantaloon cast test (PCT), Modic changes, and a summary of physical symptoms, coined "loading factor." By means of multivariate stepwise regression, prognostic factors that were useful in predicting outcomes were identified.

RESULTS: A total of 91 patients fit the inclusion criteria, with a mean follow-up of 33±16 months. Discography, Modic changes, and loading factor were of no value for predicting outcome scores ($p>.05$). A positive PCT predicted improved outcomes in back pain severity, but only in patients without prior surgery ($p=.02$). Demographic factors that showed a consistent reduction in back pain were female sex ($p=.021$) and no prior surgery at index level ($p=.009$). No other sociodemographic factors were of predictive value ($p>.05$).

CONCLUSIONS: In patients without prior surgery, the PCT appears to be the most promising prognostic tool. Other prognostic selection tools such as discography and Modic changes yield disappointing results. In this study, female patients and those without prior spine surgery appear to be most likely to benefit from fusion surgery for DDD.

Keywords: Disc degeneration; Discography; Lumbar fusion; Modic; Pantaloon cast; Patient selection

FDA device/drug status: Not applicable.

Author disclosures: ***VES:*** Nothing to disclose. ***PPAV:*** Nothing to disclose. ***DJZ:*** Nothing to disclose. ***MLS:*** Consulting: Mazor Robotics, Ltd (B); Speaking and/or Teaching Arrangements: Mazor Robotics, Ltd (B); Trips/Travel: Mazor Robotics, Ltd (B), outside the submitted work.

The disclosure key can be found on the Table of Contents and at www.TheSpineJournalOnline.com.

VS, PV, and DZ declare no potential conflicts of interest. MS is a consultant to Mazor Robotics, Ltd.

* Corresponding author. Department of Neurosurgery, c/o Bergman Clinics, Rijksweg 69, 1411 GE Naarden, The Netherlands. Tel.: 0031 88 900 0500; fax: 0031 88 900 0568.

E-mail address: victor.staartjes@gmail.com (V.E. Staartjes)

https://doi.org/10.1016/j.spinee.2017.08.242
1529-9430/

Introduction

Low back pain is one of the top-three causes of disability in Western societies and imposes significant direct and indirect socioeconomic costs [1,2]. The etiology of low back pain is multifactorial, but it is often related to degenerative disc disease (DDD) [3–5]. The standard treatment for progressive DDD in patients who are unresponsive to long-term conservative treatment is interbody fusion, but this is controversial [6]. With some reports showing no benefit compared with conservative treatment, patient selection is vitally important [7,8]. Various prognostic tests attempt to identify subsets of patients that might benefit most from surgery, but the validity of these tests is unclear [7].

The most widely used preoperative test is provocative discography [9], which aims to determine to what extent the disc is responsible for symptoms [7]. Because weak evidence exists for the usefulness of discography, and cytotoxic and proinflammatory sequelae of the injection may accelerate degeneration, examining the risk-benefit ratio of this invasive method is crucial [10–14].

Discogenic pain is assumed to originate from nerve ingrowth into the innermost disc mediated by proinflammatory cytokines, for example, interleukin-1, tumor necrosis factor-α, and nerve growth factor [15,16]. The expression of these cytokines is presumed to be amplified by hypermobility, which is caused by dehydration and disc degradation by matrix metalloproteinases, for example, MMP1, ADAMTS-4, and ADAMTS-5 [16,17]. Modic-type end plate changes, which are characterized by edematous (type I), fatty (II), or sclerotic (III) turnover, are identified using magnetic resonance imaging (MRI) [18–21]. These changes could play a role in effective patient selection for surgery, although there is little evidence regarding the prognostic value of Modic changes [19,22,23].

Lumbar fusion aims to reduce back pain by stabilizing and relieving degenerated segments. The pantaloon cast test (PCT) and external transpedicular fixation aim to simulate spinal fusion by limiting lumbar joint mobility [24,25]. If successful in alleviating pain, it is assumed that fusion will yield favorable results. Similarly, physical symptoms can provide clues about whether stabilizing the segment will improve symptoms. Recently, the daily course of complaints and the influence of rest, mobility, and posture have been identified as relevant indicators [26,27], which we have integrated as "loading factor" (LF).

Ultimately, success in this category of patients should be defined by improved physical symptoms (patient-reported outcome measures) rather than technical success of the procedure. The aim of this cohort study was to assess if prognostic tests and preoperative sociodemographic factors are useful for the outcome-oriented selection of DDD patients for fusion surgery.

Materials and methods

Overview

We reviewed the prospectively recorded data of all patients with DDD who underwent single-level lumbar interbody fusion at a single center (2010–2016). The preoperative tests used were discography and PCT. The LF was assessed during outpatient clinics. Modic changes were assessed using MRI (Magnetom Essenza, Siemens, Munich, Germany, 1.5 Tesla). Patients were followed-up at 6 weeks, 12 months, and 24 months postoperatively. In June 2016, there was a final mailed follow-up for visual analog scale (VAS) scores for back (VAS-BP) and leg pain (VAS-LP), Oswestry Disability Index (ODI), sociodemographic data, and whether further treatment was received [28]. This study was approved by the Dutch Central Committee on Research Involving Human Subjects. Informed consent was obtained from all individual participants included in this study.

Study population and patient selection

The study inclusion criteria were a complete preoperative record, a minimum follow-up of 12 months, and DDD diagnosed by MRI. An a priori sample size calculation using SPSS SamplePower 3 (SPSS, Chicago, IL, USA) concluded that for regression analysis with three predictors, anticipated f^2=0.15, and power of 0.8, a minimum of 76 patients would be needed.

The preoperative exclusion criteria were body mass index (BMI) ≥33, age >80, and multilevel disc pathology. Surgical patient selection was performed in a strict fashion. Prognostic tests were only considered if patients had experienced at least 6 months of severe intractable low back pain and undergone various conservative treatments. Discography and PCT were then performed whenever feasible, and routinely formed the basis for surgical decision making. If both tests were negative, patients were not considered for surgery. In all other cases, surgery was recommended as a "last resort." Smokers were strongly advised to quit smoking before surgery.

Discography

All invasive tests were performed by the same anesthetist using consistent technique. Fluoroscopy-guided provocative discography with 1–2 mL of Xenetix 300 (Guerbet, Villepinte, France) was followed by discoblock with injection of 2–3 mL of lidocaine (2%) into the center of the disc. For a positive outcome, a clear pain response upon provocation and a significant alleviation of low back pain (subjective ≥50% pain reduction) after discoblock was necessary. Otherwise, the test was counted as negative.

Pantaloon cast test

Patients were placed in a plaster cast with a pantaloon shape (Fig. 1) covering the lumbar region up to around T10 and one

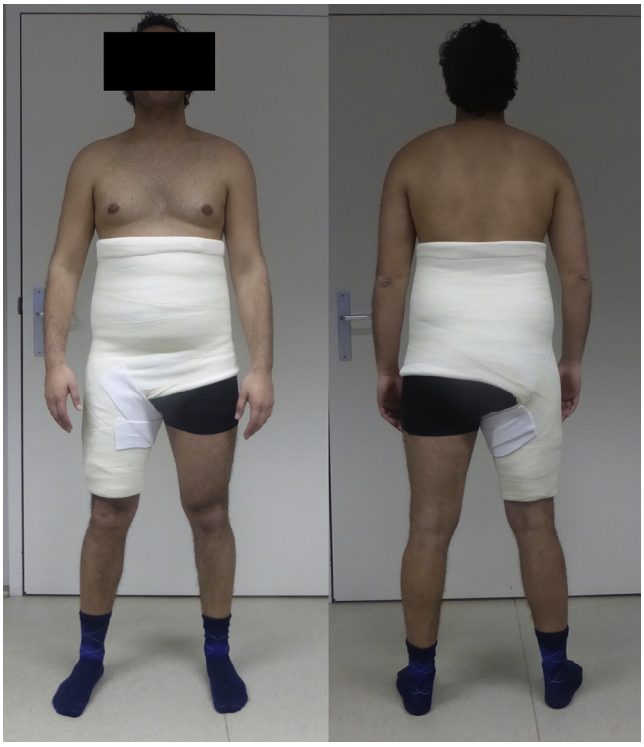

Fig. 1. A hard plaster cast (pantaloon cast) used for preoperative testing, covering, and stabilizing the lumbar spine.

thigh of the patient's choice in an attempt to simulate fixation. All casts were fitted by the same plaster technician. Patients were asked to wear the cast for as long as feasible (up to 14 days), continuing even after significant pain improvement had already been achieved, and instructed to move as freely as possible. A pain reduction of ≥50% was counted as a positive result.

Loading factor

This novel factor was determined from patient history. If patients experienced significantly more back pain during the day or during axial loading (eg, exercise, prolonged sitting, or standing), LF was counted as positive (LF+). The LF was described as negative if patients experienced more pain during the night or in supine position (LF–). If patients felt that activity did not make a notable difference, LF was recorded as neutral (LF0).

Surgical techniques

In the presence of additional radicular symptoms, minimally invasive transforaminal lumbar interbody fusion was the surgical technique of choice, and was performed using the standard technique [29] with pedicle screws inserted percutaneously using robotic guidance (SpineAssist, Mazor Robotics Ltd, Caesarea, Israel). In the absence of radicular symptoms, we opted for transaxial lumbosacral fusion (AxiaLIF, TranS1, Denver, CO, USA) or anterior lumbar interbody fusion (ALIF). Especially in young men, AxiaLIF was chosen over ALIF because of the latter's inherent risk of retrograde ejaculation. AxiaLIF was carried out as previously described [30], and ALIF was performed using a mini-open approach according to the technique described by Brau [31] using a screw-augmented cage (Synfix, Synthes Spine, Inc, West Chester, PA, USA).

Data and statistics

Continuous variables were reported as means±standard deviations and categorical variables as percentages. Clinical success, and thus the minimal clinically important difference (MCID), was defined as ≥30% improvement [32]. We stratified continuous variables into two groups for dichotomous analysis (BMI ≤25:>25 kg/m^2, age ≤40:>40 years, history of pain ≤12:>12 months).

Data were analyzed using SPSS V24.0 (IBM SPSS, IBM Corp, Armonk, NY, USA). Wilcoxon tests were used to compare related data, and Bonferroni correction for multiple testing was applied. For 2×2 comparison of categorical data, a chi-square was used. Prognostic variables for each outcome score were analyzed using Kruskal-Wallis tests. The cutoff for further analysis was p≤.2 to achieve broad inclusion. The remaining variables were entered into a multivariate stepwise linear regression model to predict change scores for VAS and ODI. The least significant variable was repeatedly excluded until only significant prognostic factors remained. The same procedure, but using logistic regression, was performed to identify prognostic factors for MCID. A two-tailed p≤.05 was regarded as significant.

Results

During the study period, 124 patients with DDD, all of whom had undergone at least 6 months of unsuccessful conservative treatment, were observed. Fig. 2 shows a flowchart of the study population. After appropriate clinical selection, 91 patients (73.4%) underwent surgery; all met the inclusion criteria. Thirty-three patients (26.6%) were not operated as they either had two negative tests or had not undergone sufficient conservative treatment. Of the surgical patients, 72 (79.1%) underwent discography, whereas 65 (71.4%) underwent the PCT. The LF and Modic changes were identified in all cases. No dropouts were recorded. Detailed baseline patient characteristics are shown in Table 1.

Surgical technique did not influence outcomes (all p>.2). At baseline, 53 patients (58.2%) had not undergone prior surgery at index level. Perioperative and outcome data are summarized in Table 2. Compared with baseline, VAS-BP improved by 47.5±30.6 points (–57.8%), VAS-LP improved by 21.8±40.6 (–48.0%), and ODI improved by 29.8±21.0 (–59.0%) (all p<.001) at a final follow-up (33±16 months). MCID was achieved in 65 patients (71.4%) for VAS-BP, 47 (51.6%) for VAS-LP, and 69 (75.8%) for ODI. Ability to work increased from 57.3% to 80.8% (p<.001).

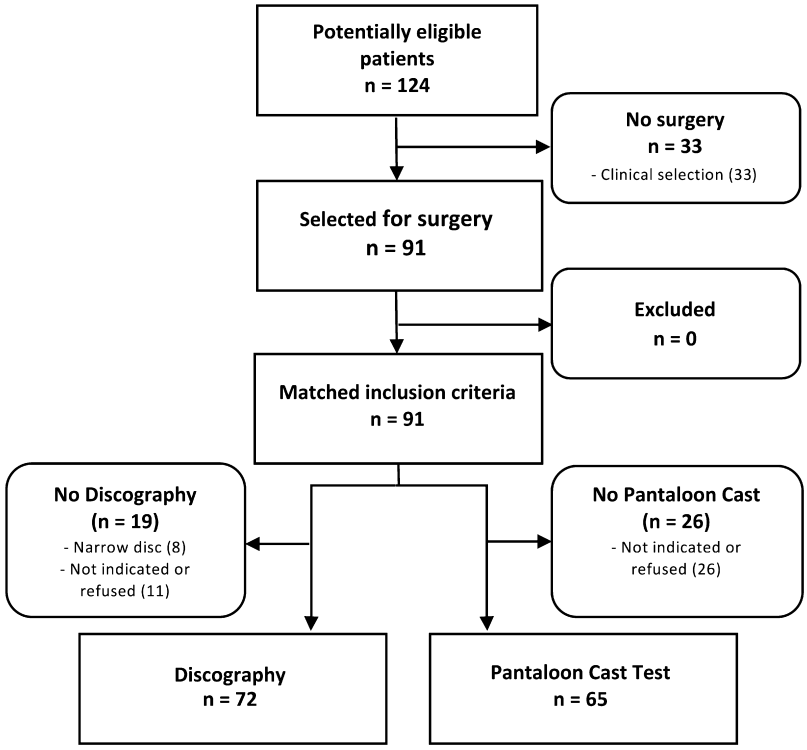

Fig. 2. Study flow chart.

Analgesic medication was fully discontinued by 57.5% at the final follow-up, whereas 19.2% still used these medications daily. In total, 61.6% of patients were satisfied, and 90.4% would choose this treatment again.

One (1.1%) infection and one (1.1%) permanent quadriceps paresis were encountered. Revision surgery was performed for persisting pain in five (5.5%) cases, and a Tarlov cyst (index level, achieved MCID) in one (1.1%) case. Table 3 summarizes the outcomes of the predictive tests. Fig. 3 provides boxplots of the outcomes. The PCT was worn for a mean of 12.7±2.7 days (range 4–15 days), with 46 patients (70.8%) reaching the 14-day threshold. No adverse events were observed during testing. Older patients were more likely to have a positive PCT outcome (p=.002), whereas BMI and sex had no influence (both p>.05).

Table 1
Baseline population statistics

Baseline	No.	Range	Standard deviation
Male sex	37 (40.7%)		
No prior surgery at index level	53 (58.2%)		
Index level			
L2–L3	1 (1.1%)		
L3–L4	2 (2.2%)		
L4–L5	22 (24.2%)		
L5–S1	66 (72.5%)		
Smoking status			
Smoker	30 (33%)		
Quit smoking	20 (21.6%)		
Non-smoker	38 (41.7%)		
Preoperative ability to work			
Fully able	23 (25.3%)		
Limited	28 (30.8%)		
Not able	38 (41.7%)		
Age (y)	43	19–64	10.4
Body mass index (kg/m^2)	24.5	17.9–32.7	3
History of back and leg pain (mo)	21.6	4–120	29.5
Back pain severity (VAS)	82.2	40–100	12.6
Leg pain severity (VAS)	45.4	0–100	30.5
Oswestry Disability Index	50.5	16–84	13.7

VAS, visual analog scale.

Table 2
Perioperative data and general clinical outcomes

Characteristic	No.	Range	Standard deviation
Length of follow-up (mo)	33.3	12–60	15.8
Surgical technique			
AxiaLIF	47 (51.6%)		
ALIF	23 (25.3%)		
MI-TLIF	21 (23.1%)		
Perioperative data			
Estimated blood loss (mL)	97.0	50–2,500	184.9
Dose area product (cGy cm^2)	314.6	44.9–818.6	162.5
Length of surgery (min)	92.6	36–274	64.3
Length of stay (d)	1.5	0–5	0.7
Clinical outcomes at final follow-up			
Back pain severity (VAS)	34.7	0–90	29.5
Leg pain severity (VAS)	23.6	0–90	29.0
Oswestry Disability Index	20.7	0–68	19.3

AxiaLIF, transaxial lumbar interbody fusion; ALIF, anterior lumbar interbody fusion; MI-TLIF, minimally invasive transforaminal interbody fusion; VAS, visual analog scale.

Table 3
Results of preoperative predictive tests (p-values for the mean differences are provided)

Test	Result	VAS-BP change score	p	VAS-LP change score	p	ODI change score	p
Discography							
Positive	63 (69.2%)	50.2±29.4	.41	27.7±39.8	.04*	29.6±20.7	.39
Negative	9 (9.9%)	40.6±32.1		5.0±33.9		23.0±22.7	
Pantaloon cast test							
Positive	55 (60.4%)	52.2±27.8	.15	22.9±37.5	.63	30.5±21.3	.99
Negative	10 (11.0%)	34.0±34.3		28.1±48.5		27.3±22.5	
Modic changes							
None	19 (20.9%)	49.0±28.1	.65	23.2±48.4	.22	29.6±22.1	.95
Type 1	31 (34.1%)	42.9±33.1		16.1±36.3		29.8±20.5	
Type 2	41 (45.1%)	50.2±30.0		30.4±40.5		28.6±21.3	
Loading factor							
+	44 (48.4%)	50.5±28.2	.35	22.6±38.7	.58	30.4±20.2	.30
0	23 (25.3%)	40.0±31.8		23.0±45.2		26.8±20.4	
–	9 (9.9%)	40.0±28.3		11.7±30.8		21.6±19.4	

VAS-BP, visual analog scale for back pain; VAS-LP, visual analog scale for leg pain; ODI, Oswestry Disability Index.
* p≤.05.

Univariate analysis

Univariate analysis was performed for surgical technique, discography, PCT, Modic changes, LF, prior surgery, age, sex, BMI, index level, preoperative smoking status, preoperative ability to work, history of pain in months, and occupation. Table 4 lists all prognostic factors with p≤.2.

Multivariate linear regression analysis

No prior surgery at index level (p=.009) and female sex (p=.021) were predictors of back pain improvement, explaining 12.8% of the variance. No prior surgery (p=.002) was also a predictor for achievement of MCID in leg pain, explaining 14.4% of the variance. There was no difference in the

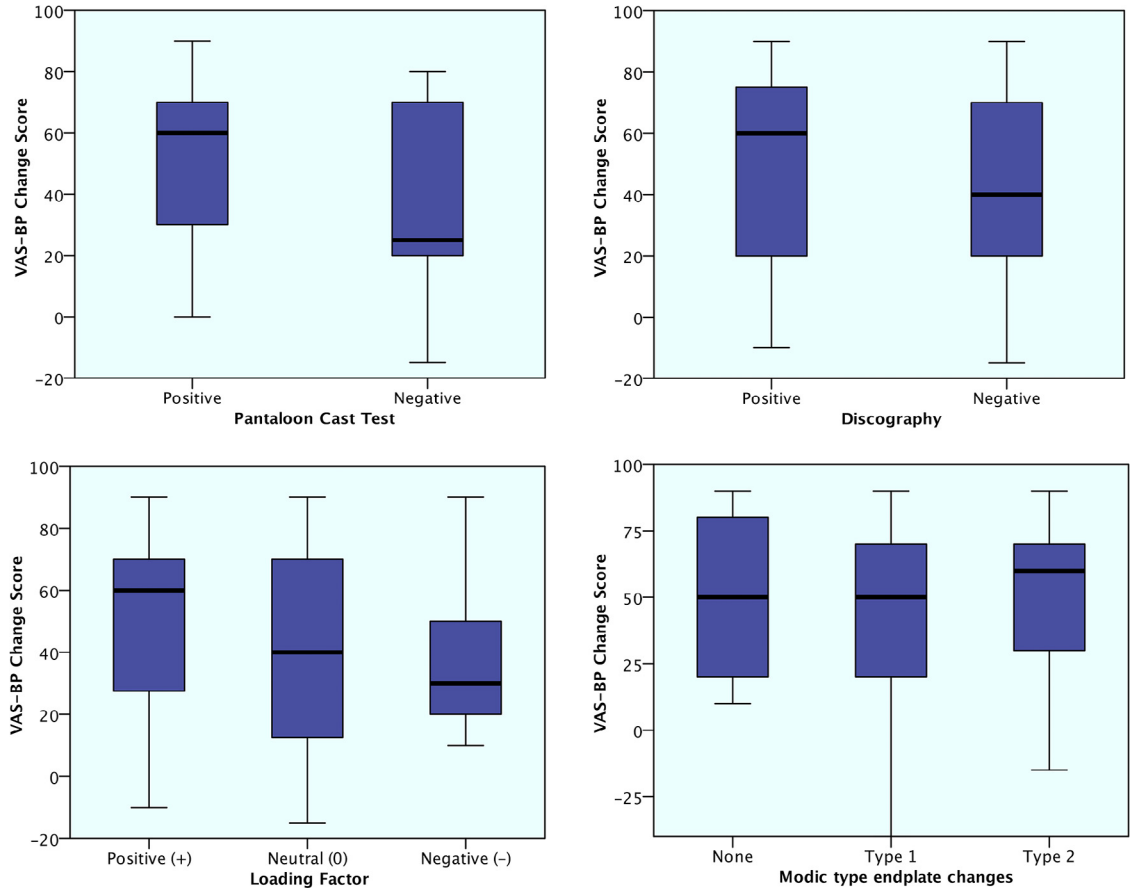

Fig. 3. Boxplots showing the effect of prognostic factors on the change score for back pain severity. VAS-BP, visual analog scale for back pain.

Table 4
Univariate analysis of prognostic factors for VAS and ODI

Prognostic factor	VAS-BP		VAS-LP		ODI	
	p	η^2	p	η^2	p	η^2
Ability to work	.107	0.06				
Discography			.043	0.07		
Index level					.126	0.06
Pantaloon cast test	.147	0.03				
Primary discopathy	.012	0.07	.152	0.03	.192	0.02
Sex	.036	0.05			.059	0.04
Type of work			.196	0.10		

VAS-BP, visual analog scale for back pain; VAS-LP, visual analog scale for leg pain; ODI, Oswestry Disability Index.

Only prognostic factors with p≤.2 in the univariate analysis are represented. The eta-squared effect size is reported as η^2.

number of prior surgeries between sexes (male: 16 of 37 (43.3%), female: 22 of 54 (40.7%), p=.81), strengthening the evidence of these being independent predictors. No other significant predictors were found. When performing the same analysis exclusively for primary discopathies, only positive PCT (p=.02) was a significant predictor for back pain improvement, explaining 12.3% of the variance. Table 5 shows the results of regression analysis. Male sex (p=.013) was the only predictor of unimproved or worsening back pain.

Discussion

No prior surgery and female sex were significantly associated with improved long-term outcomes after fusion, whereas a positive PCT result was only helpful in identifying better fusion candidates in a subgroup of primary discopathy patients. Other commonly used prognostic factors, such as discography, Modic changes, and LF, showed little to no value.

Discography was not a useful predictor of outcomes in our cohort. This is in line with reviews, which all failed to show a significant prognostic effect [7,30,33,34]. However, randomized controlled trials (RCTs) by Margetic et al. [12] and Lee et al. [35] found a positive predictive effect. We used a combination of provocative discography and discoblock. An RCT by Ohtori et al. [36] showed the superiority of discoblock over provocative discography. Because there is a striking lack of uniformity among discography techniques (ie, use of provocation or discoblock, and injection pressures), it is hard to draw firm conclusions. Overall, there is little supporting evidence for the use of traditional provocative discography [9], and it should therefore not be used in routine clinical practice, especially because it can cause clinically relevant injury to the disc [13,14]. We have ceased using discography as a tool for patient selection, but it is still used throughout Europe. That positive discography predicted a greater reduction in leg pain (p=.04) is remarkable, and could be a consequence of numbing the branches of the communicating rami that innervate the lateral disc, thus eliminating pseudoradicular pain.

The PCT did not significantly predict outcome scores in the whole cohort. However, we analyzed a subgroup of 53 patients without prior surgery and found a significant prognostic effect (p=.02). This correlates with a review by Willems et al. [25], which also found the PCT of value for patients without prior surgery. In addition, it was found that older patients were more likely to have a positive PCT outcome.

Although there is little data available on the PCT, owing to its non-invasiveness, it is a viable option when dealing with primary discopathy in patients willing and able to tolerate this cumbersome test. We have routinely used the PCT in clinical practice for 13 years and acknowledge that its cumbersomeness could be a confounder. Patients who desire surgery may more easily report a positive test result (Hawthorne effect) [37].

Nonetheless, DDD patients who benefit most from fusion surgery are those who have suffered from intractable pain for years, tried many conservative treatment strategies in vain, and are almost desperate for surgery as a "last resort." The PCT assesses not only physical components as a "simulation" of fusion, but perhaps also the patients' mind-set and their determination for a good outcome.

Modic et al. [18] described tissue changes that are strongly related to low back pain. There is an ongoing debate regarding the etiology of these changes [21,22,38–40]. Although

Table 5
Stepwise multiple regression models

Model	Factors included in model	Means	p	R^2
VAS-BP change score (Full cohort)	Primary discopathy No prior surgery-to-prior surgery ratio n=53:n=38	54.4±29.1:37.8±30.3	.009	0.128 (p=.002)
	Sex Female-to-male ratio n=54:n=37	53.5±27.2:38.6±33.4	.021	
MCID VAS-LP (Full cohort)	Primary discopathy No prior surgery-to-prior surgery ratio n=53:n=38	81.1%:50% achieved MCID	.002	0.144 (p=.002)
VAS-BP change score (Primary cases)	Pantaloon cast test Positive-to-negative ratio n=37:n=7	56.9±26.8:29.3±27.8	.02	0.123 (p=.02)

VAS-BP, visual analog scale for back pain; MCID VAS-LP, minimal clinically important difference in visual analog scale for leg pain (≥30%).
Only models with significant prognostic factors are represented. The p-value of R^2 represents the significance of the model.

these are considered a promising prognostic factor, Modic changes were not clinically predictive in this study. Laustsen et al. [19] and Ohtori et al. [41] found similar results. Kwon et al. [42] only found that the rare type III was related to worse outcomes. It currently seems unlikely that Modic changes have a generalizable and clinically relevant prognostic value for fusion surgery.

Patients with LF+ showed greater improvement than those with LF– or LF0. A clear trend was observed, albeit without statistical significance (Fig. 3). The advantage of LF is that it encapsulates many aspects of physical symptoms that can be clinically relevant: pain interference in daily activities and walking, the diurnal course of complaints, and the influence of rest, mobility, and posture. Their predictive value for patient selection in spine surgery is controversial [26,27]. Our findings did not provide a conclusive result, potentially due to the small sample size.

Patients without previous surgery had significantly better outcomes for back pain. This is not surprising, because surgery for failed back surgery syndrome is less predictable [26,43–45]. Female sex was a strong positive prognostic factor, which is likely due to female patients often presenting with worse complaints preoperatively, coupled with greater postoperative improvement compared with male patients [26,30,46]. Furthermore, male sex was the only predictor of worsening back pain, thus supporting this finding. Body mass index had no effect in this cohort, but this is probably due to the exclusion of severely overweight patients from surgery. Although an effect of obesity on outcomes seems plausible and has been demonstrated in the Spine Patient Outcomes Research Trial (SPORT) trial [47], this finding is supported by multiple studies [26,48,49]. Consistent with the literature, age was also of no predictive value [26,48,50].

All recent RCTs show that fusion surgery does not produce significantly better results than conservative treatment [8,51]. This is especially true because physical symptoms can be due to pseudoradicular nerve or facet joint pathology rather than disc degeneration seen on MRI. Although surgery may not provide more success than conservative treatment for DDD in the general patient population, there are subsets of patients that will truly benefit more from fusion surgery. Rigorous patient selection is key to success, and although finding these subsets is exceedingly difficult as there are no robust selection tools available, there are some promising prognostic factors that may guide spine surgeons in this respect. Traditionally, discography has seen wide adoption, but has repeatedly failed to prove useful in many trials and may even do more harm than good [7,9,12–14,30,33–36]. On this basis, we cannot recommend the use of discography for surgical patient selection and have ceased using it routinely.

Although cumbersome, the PCT has shown a prognostic effect in patients without previous surgery, as confirmed by another study [25]. We feel that there is a lack of literature on this test because many see it as archaic, but it may be the best prognostic tool available. For this reason, we have recently initiated an RCT to systematically trial the PCT on a higher level of evidence.

We analyzed multiple clinically relevant variables in a homogenous cohort of single-level discopathy patients. In addition, we provided one of the longest follow-ups for discopathy after fusion surgery.

We used comparatively strict statistical analyses to compensate for cohort size. Although the cohort was small, an a priori power analysis concluded that the minimum power of 0.8 was reached. Some subgroup analyses in this study did not reach the required minimum sample size of 76 patients and may be underpowered. Post hoc power analyses indicated that the analyses of the PCT subgroup (65 patients), and the primary discopathy subgroup (53 patients) reached statistical powers of 0.798 and 0.697, respectively. Although our dataset is relatively complete for a retrospective study, it would have been interesting to follow up patients who were not treated surgically. The number of negative PCT and discography tests was relatively small, and not all patients underwent both discography and PCT. Finally, one author (MS) had a potential conflict of interest as a consultant to Mazor Robotics, Ltd, although this is not directly related to this study.

Conclusions

Patient selection for fusion surgery for DDD remains difficult but is key to clinical success. Female patients without prior surgery at index level are most likely to benefit from fusion surgery. Preoperative selection tools such as discography and Modic changes repeatedly yield disappointing results in clinical trials. Based on previous reports and the results of this study, we cannot recommend provocative discography for surgical patient selection. In a subset of patients without prior surgery at index level, the PCT helps in surgical decision making. We do not recommend the use of these tests as absolute “red light” or “green light” indicators, but advocate carefully balancing all other clinical data against test results and patient expectations. Rather, their prognostic value should be conferred to the patient to help create a realistic expectation pattern. The same applies to sociodemographic prognostic factors such as gender, which should not form the basis of surgical decision making.

References

[1] Lambeek LC, van Tulder MW, Swinkels ICS, Koppes LLJ, Anema JR, van Mechelen W. The trend in total cost of back pain in The Netherlands in the period 2002 to 2007. Spine 2011;36:1050–8. doi:10.1097/BRS.0b013e3181e70488.

[2] GBD 2013 DALYs and HALE Collaborators, Murray CJL, Barber RM, Foreman KJ, Abbasoglu Ozgoren A, Abd-Allah F, et al. Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990–2013: quantifying the epidemiological transition. Lancet 2015;386:2145–91. doi:10.1016/S0140-6736(15)61340-X.

[3] Cheung KMC, Karppinen J, Chan D, Ho DWH, Song Y-Q, Sham P, et al. Prevalence and pattern of lumbar magnetic resonance imaging changes in a population study of one thousand forty-three individuals. Spine 2009;34:934–40. doi:10.1097/BRS.0b013e3181a01b3f.
[4] Wang Z-X, Hu Y-G. High-intensity zone (HIZ) of lumbar intervertebral disc on T2-weighted magnetic resonance images: spatial distribution, and correlation of distribution with low back pain (LBP). Eur Spine J 2012;21:1311–15. doi:10.1007/s00586-012-2240-0.
[5] Jacobs JJ, Andersson GB, Bell J-E, Weinstein SL, Dormans JP, Gnatz SM, et al., editors. Spine: low back and neck pain. Rosemont, IL: United State Bone and Joint Initiative: The Burden of Musculoskeletal Disease in the United States, American Academy of Orthopaedic Surgeons; 2011. p. 21–56.
[6] Manchikanti L, Falco FJE, Singh V, Benyamin RM, Racz GB, Helm S, et al. An update of comprehensive evidence-based guidelines for interventional techniques in chronic spinal pain. Part I: introduction and general considerations. Pain Physician 2013;16:S1–48.
[7] Willems P. Decision making in surgical treatment of chronic low back pain: the performance of prognostic tests to select patients for lumbar spinal fusion. Acta Orthop Suppl 2013;84:1–35. doi:10.3109/17453674.2012.753565.
[8] Mannion AF, Brox J-I, Fairbank JC. Consensus at last! Long-term results of all randomized controlled trials show that fusion is no better than non-operative care in improving pain and disability in chronic low back pain. Spine J 2016;16:588–90. doi:10.1016/j.spinee.2015.12.001.
[9] Eck JC, Sharan A, Resnick DK, Watters WC, Ghogawala Z, Dailey AT, et al. Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 6: discography for patient selection. J Neurosurg Spine 2014;21:37–41. doi:10.3171/2014.4.SPINE14269.
[10] Putzier M, Streitparth F, Hartwig T, Perka CF, Hoff EK, Strube P. Can discoblock replace discography for identifying painful degenerated discs? Eur J Radiol 2013;82:1463–70. doi:10.1016/j.ejrad.2013.03.022.
[11] Manchikanti L, Glaser SE, Wolfer L, Derby R, Cohen SP. Systematic review of lumbar discography as a diagnostic test for chronic low back pain. Pain Physician 2009;12:541–59.
[12] Margetic P, Pavic R, Stancic MF. Provocative discography screening improves surgical outcome. Wien Klin Wochenschr 2013;125:600–10. doi:10.1007/s00508-013-0404-5.
[13] Cuellar JM, Stauff MP, Herzog RJ, Carrino JA, Baker GA, Carragee EJ. Does provocative discography cause clinically important injury to the lumbar intervertebral disc? A 10-year matched cohort study. Spine J 2016;16:273–80. doi:10.1016/j.spinee.2015.06.051.
[14] Carragee EJ, Don AS, Hurwitz EL, Cuellar JM, Carrino JA, Carrino J, et al. 2009 ISSLS Prize Winner: does discography cause accelerated progression of degeneration changes in the lumbar disc: a ten-year matched cohort study. Spine 2009;34:2338–45. doi:10.1097/BRS.0b013e3181ab5432.
[15] Ohtori S, Inoue G, Miyagi M, Takahashi K. Pathomechanisms of discogenic low back pain in humans and animal models. Spine J 2015;15:1347–55. doi:10.1016/j.spinee.2013.07.490.
[16] Vergroesen P-PA, Kingma I, Emanuel KS, Hoogendoorn RJW, Welting TJ, van Royen BJ, et al. Mechanics and biology in intervertebral disc degeneration: a vicious circle. Osteoarthr Cartil 2015;23:1057–70. doi:10.1016/j.joca.2015.03.028.
[17] Landi A, Gregori F, Mancarella C, Maiola V, Maccari E, Marotta N, et al. Lumbar spinal degenerative "microinstability": hype or hope? Proposal of a new classification to detect it and to assess surgical treatment. Eur Spine J 2015;24(Suppl. 7):872–8. doi:10.1007/s00586-015-4274-6.
[18] Modic MT, Steinberg PM, Ross JS, Masaryk TJ, Carter JR. Degenerative disk disease: assessment of changes in vertebral body marrow with MR imaging. Radiology 1988;166:193–9. doi:10.1148/radiology.166.1.3336678.
[19] Laustsen AF, Bech-Azeddine R. Do Modic changes have an impact on clinical outcome in lumbar spine surgery? A systematic literature review. Eur Spine J 2016;25:3735–45. doi:10.1007/s00586-016-4609-y.
[20] Mok FPS, Samartzis D, Karppinen J, Fong DYT, Luk KDK, Cheung KMC. Modic changes of the lumbar spine: prevalence, risk factors, and association with disc degeneration and low back pain in a large-scale population-based cohort. Spine J 2016;16:32–41. doi:10.1016/j.spinee.2015.09.060.
[21] Dudli S, Fields AJ, Samartzis D, Karppinen J, Lotz JC. Pathobiology of Modic changes. Eur Spine J 2016;25:3723–34. doi:10.1007/s00586-016-4459-7.
[22] Zhang Y-H, Zhao C-Q, Jiang L-S, Chen X-D, Dai L-Y. Modic changes: a systematic review of the literature. Eur Spine J 2008;17:1289–99. doi:10.1007/s00586-008-0758-y.
[23] Nguyen C, Poiraudeau S, Rannou F. From Modic 1 vertebral-endplate subchondral bone signal changes detected by MRI to the concept of "active discopathy. Ann Rheum Dis 2015;74:1488–94. doi:10.1136/annrheumdis-2015-207317.
[24] Rask B, Dall BE. Use of the pantaloon cast for the selection of fusion candidates in the treatment of chronic low back pain. Clin Orthop Relat Res 1993;148–57.
[25] Willems PC, Elmans L, Anderson PG, Jacobs WCH, van der Schaaf DB, de Kleuver M. The value of a pantaloon cast test in surgical decision making for chronic low back pain patients: a systematic review of the literature supplemented with a prospective cohort study. Eur Spine J 2006;15:1487–94. doi:10.1007/s00586-006-0121-0.
[26] van Hooff ML, van Loon J, van Limbeek J, de Kleuver M. The Nijmegen decision tool for chronic low back pain. Development of a clinical decision tool for secondary or tertiary spine care specialists. PLoS ONE 2014;9:doi:10.1371/journal.pone.0104226. e104226.
[27] Ohtori S, Orita S, Yamauchi K, Eguchi Y, Aoki Y, Nakamura J, et al. Do physical symptoms predict the outcome of surgical fusion in patients with discogenic low back pain? Asian Spine J 2016;10:509–15. doi:10.4184/asj.2016.10.3.509.
[28] Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiotherapy 1980;66:271–3.
[29] Peng CWB, Yue WM, Poh SY, Yeo W, Tan SB. Clinical and radiological outcomes of minimally invasive versus open transforaminal lumbar interbody fusion. Spine 2009;34:1385–9. doi:10.1097/BRS.0b013e3181a4e3be.
[30] Zeilstra DJ, Staartjes VE, Schröder ML. Minimally invasive transaxial lumbosacral interbody fusion: a ten year single-centre experience. Int Orthop 2017;41:113–19. doi:10.1007/s00264-016-3273-5.
[31] Brau SA. Mini-open approach to the spine for anterior lumbar interbody fusion: description of the procedure, results and complications. Spine J 2002;2:216–23.
[32] Ostelo RWJG, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine 2008;33:90–4. doi:10.1097/BRS.0b013e31815e3a10.
[33] Willems PC, Elmans L, Anderson PG, van der Schaaf DB, de Kleuver M. Provocative discography and lumbar fusion: is preoperative assessment of adjacent discs useful? Spine 2007;32:1094–9. doi:10.1097/01.brs.0000261672.97430.b0. discussion 1100.
[34] Carragee EJ, Lincoln T, Parmar VS, Alamin T. A gold standard evaluation of the "discogenic pain" diagnosis as determined by provocative discography. Spine 2006;31:2115–23. doi:10.1097/01.brs.0000231436.30262.dd.
[35] Lee CK, Shin DA, Kim HI, Yi S, Ha Y, Kim KN, et al. Automated pressure-controlled discography in patients undergoing anterior lumbar interbody fusion for discogenic back pain. World Neurosurg 2017;97:8–15. doi:10.1016/j.wneu.2016.09.019.
[36] Ohtori S, Kinoshita T, Yamashita M, Inoue G, Yamauchi K, Koshi T, et al. Results of surgery for discogenic low back pain: a randomized study using discography versus discoblock for diagnosis. Spine 2009;34:1345–8. doi:10.1097/BRS.0b013e3181a401bf.
[37] Adair JG. The Hawthorne effect: a reconsideration of the methodological artifact. J Appl Psychol 1984;69:334–45. doi:10.1037/0021-9010.69.2.334.

[38] Artus M, Fairbank J, Scarborough M, Foster N. Infection and low back pain: seeking evidence or fear of exploring new indications for antibiotics? Eur Spine J 2016;25:3859–61. doi:10.1007/s00586-016-4513-5.

[39] Kerttula L, Luoma K, Vehmas T, Grönblad M, Kääpä E. Modic type I change may predict rapid progressive, deforming disc degeneration: a prospective 1-year follow-up study. Eur Spine J 2012;21:1135–42. doi:10.1007/s00586-012-2147-9.

[40] Luoma K, Vehmas T, Kerttula L, Grönblad M, Rinne E. Chronic low back pain in relation to Modic changes, bony endplate lesions, and disc degeneration in a prospective MRI study. Eur Spine J 2016;25:2873–81. doi:10.1007/s00586-016-4715-x.

[41] Ohtori S, Yamashita M, Yamauchi K, Inoue G, Koshi T, Suzuki M, et al. Change in Modic type 1 and 2 signals after posterolateral fusion surgery. Spine 2010;35:1231–5. doi:10.1097/BRS.0b013e3181bde562.

[42] Kwon Y-M, Chin D-K, Jin B-H, Kim K-S, Cho Y-E, Kuh S-U. Long term efficacy of posterior lumbar interbody fusion with standard cages alone in lumbar disc diseases combined with Modic changes. J Korean Neurosurg Soc 2009;46:322–7. doi:10.3340/jkns.2009.46.4.322.

[43] Arts MP, Kols NI, Onderwater SM, Peul WC. Clinical outcome of instrumented fusion for the treatment of failed back surgery syndrome: a case series of 100 patients. Acta Neurochir (Wien) 2012;154:1213–17. doi:10.1007/s00701-012-1380-7.

[44] Gillard DM, Corenman DS, Dornan GJ. Failed less invasive lumbar spine surgery as a predictor of subsequent fusion outcomes. Int Orthop 2014;38:811–15. doi:10.1007/s00264-013-2167-z.

[45] Jenkins LT, Jones AL, Harms JJ. Prognostic factors in lumbar spinal fusion. Contemp Orthop 1994;29:173–80.

[46] Triebel J, Snellman G, Sandén B, Strömqvist F, Robinson Y. Women do not fare worse than men after lumbar fusion surgery: two-year follow-up results from 4,780 prospectively collected patients in the Swedish National Spine Register with lumbar degenerative disc disease and chronic low back pain. Spine J 2016;doi:10.1016/j.spinee.2016.11.001.

[47] Rihn JA, Radcliff K, Hilibrand AS, Anderson DT, Zhao W, Lurie J, et al. Does obesity affect outcomes of treatment for lumbar stenosis and degenerative spondylolisthesis? Analysis of the Spine Patient Outcomes Research Trial (SPORT). Spine 2012;37:1933–46. doi:10.1097/BRS.0b013e31825e21b2.

[48] Cobo Soriano J, Sendino Revuelta M, Fabregate Fuente M, Cimarra Díaz I, Martínez Ureña P, Deglné Meneases R. Predictors of outcome after decompressive lumbar surgery and instrumented posterolateral fusion. Eur Spine J 2010;19:1841–8. doi:10.1007/s00586-010-1284-2.

[49] Sielatycki JA, Chotai S, Stonko D, Wick J, Kay H, McGirt MJ, et al. Is obesity associated with worse patient-reported outcomes following lumbar surgery for degenerative conditions? Eur Spine J 2016;25:1627–33. doi:10.1007/s00586-016-4460-1.

[50] Alentado VJ, Caldwell S, Gould HP, Steinmetz MP, Benzel EC, Mroz TE. Independent predictors of a clinically significant improvement after lumbar fusion surgery. Spine J 2017;17:236–43. doi:10.1016/j.spinee.2016.09.011.

[51] Mannion AF, Brox JI, Fairbank JCT. Comparison of spinal fusion and nonoperative treatment in patients with chronic low back pain: long-term follow-up of three randomized controlled trials. Spine J 2013;13:1438–48. doi:10.1016/j.spinee.2013.06.101.