Development and validation of a prediction tool for pain reduction in adult patients undergoing elective lumbar spinal fusion: a multicentre cohort study

Esther R. C. Janssen¹ · Ilona M. Punt^{2,3} · Sander M. J. van Kuijk⁴ · Eric A. Hoebink⁵ · Nico L. U. van Meeteren⁶ · Paul C. Willems¹

Received: 17 December 2019 / Revised: 25 March 2020 / Accepted: 22 May 2020 / Published online: 29 May 2020
© The Author(s) 2020

Abstract

Purpose On average, 56% of patients report a clinically relevant reduction in pain after lumbar spinal fusion (LSF). Preoperatively identifying which patient will benefit from LSF is paramount to improve clinical decision making, expectation management and treatment selection. Therefore, this multicentre study aimed to develop and validate a clinical prediction tool for a clinically relevant reduction in pain 1 to 2 years after elective LSF.

Methods The outcomes were defined as a clinically relevant reduction in predominant (worst reported pain in back or legs) pain 1 to 2 years after LSF. Patient-reported outcome measures and patient characteristics from 202 patients were used to develop a prediction model by logistic regression. Data from 251 patients were used to validate the model.

Results Nonsmokers (odds ratio = 0.41 [95% confidence interval = 0.19–0.87]), with lower Body Mass Index (0.93 [0.85–1.01]), shorter pain duration (0.49 [0.20–1.19]), lower American Society of Anaesthesiologists score (4.82 [1.35–17.25]), higher Visual Analogue Scale score for predominant pain (1.05 [1.02–1.08]), lower Oswestry Disability Index (0.96 [0.93–1.00]) and higher RAND-36 mental component score (1.03 [0.10–1.06]) preoperatively had a higher chance of a clinically relevant reduction in predominant pain. The area under the curve of the externally validated model yielded 0.68. A nomogram was developed to aid clinical decision making.

Conclusions Using the developed nomogram surgeons can estimate the probability of achieving a clinically relevant pain reduction 1 to 2 years after LSF and consequently inform patients on expected outcomes when considering treatment.

Keywords Clinical prediction model · Decision aid · External validation · Patient-reported outcomes · Risk factors

Introduction

The number of elective lumbar spinal fusions (LSFs) has increased 2.4-fold in the past decade [1], although postoperative pain reduction often remains unsatisfactory [2]. Some patients have a considerably lower probability of achieving

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00586-020-06473-w>) contains supplementary material, which is available to authorized users.

✉ Esther R. C. Janssen
esther.janssen@mumc.nl

¹ Department of Orthopaedics and Research School Caphri, Maastricht University Medical Centre+, P. Debyelaan 25, 6229 HX Maastricht, The Netherlands

² NUTRIM School for Nutrition and Translational Research in Metabolism, Maastricht University, Maastricht, The Netherlands

³ Department of Surgery and Trauma Surgery, Maastricht University Medical Centre+, Maastricht, The Netherlands

⁴ Department of Clinical Epidemiology and Medical Technology Assessment (KEMTA), Maastricht University Medical Centre+, Maastricht, the Netherlands

⁵ Department of Orthopaedics, Amphia Hospital, Breda, The Netherlands

⁶ Top Sector Life Sciences and Health (Health~Holland), The Hague, The Netherlands

a reduction in pain postoperatively [3]. To improve clinical shared decision making, expectation management, and patient selection, it is important to predict expected outcomes after LSF and act upon this information.

Prediction tools are reliable tools that can predict the probability of outcomes after LSF. Patients and surgeons can consult such prediction tools to estimate probabilities of outcomes, such as pain reduction, after LSF for that specific patient. Factors that predict postoperative pain reduction have been reported previously [4–6]. Patient characteristics such as age, smoking, American Society of Anaesthesiologists (ASA) score and preoperative patient-reported outcome measures (PROMs) on pain, mental health and health-related quality of life (HRQOL) are associated with postoperative pain reduction [4–7]. To the best of our knowledge, only one study externally validated a prediction tool that predicts pain reduction after LSF, which has been translated into an easily implementable tool in the USA [7]. However, due to substantial differences in healthcare systems, this tool probably cannot be applied to European countries. Moreover, potentially important predictors such as symptom duration and mental health were not incorporated in that model. For use in clinical practice, an externally validated and easily applicable prediction tool developed in a representative population is imperative [8].

Thus, the aim of this multicentre cohort study is to develop and validate a prediction tool to predict the probability of clinically relevant reduction in pain 1 and 2 years after elective one- to three-level LSF.

Methods

From January 2011 until January 2015, baseline and 1- to 2-year postoperative questionnaires were collected from 202 patients undergoing elective LSF as part of routine care in the university hospital. In this cohort study, this derivation set was used to develop and internally validate the logistic regression model. The validation set was used for external validation of the model and contained baseline and 1- to 2-year postoperative data on 251 patients collected from July 2014 until November 2016 in the general hospital. This study was assessed by the local ethics committee and was considered not applicable to the Medical Research Involving Human Subject Act (number: 16-4-262.1/ivb).

Population

Adult patients (≥ 18 years) eligible for elective one- to three-level LSF were included. Diagnosis and surgical procedure were verified from their medical records. Patients were included in the study if they were diagnosed with degenerative disc disease, spondylosis, spondylolysis/-listhesis,

spinal stenosis, adjacent level disease, post-herniotomy, post-laminectomy or (recurrent) disc herniation. Revisions of a spinal fusion within 1 year of the previous surgery were excluded.

Data collection

Patients preoperatively and postoperatively completed questionnaires on the following: back and leg pain using the Visual Analogue Scale (VAS) [9], physical functioning using the Oswestry Disability Index (ODI) [10], HRQOL using the RAND-36 [11], mental health using the Pain Catastrophizing Scale (PCS) and Hospital Anxiety and Depression Scale (HADS) [12, 13]. From the three VAS scores (back pain, right leg pain and left leg pain), the predominant (worst reported) pain score was used as a predictor. The RAND-36 resulted in a mental component score (RAND-36 MCS) and a physical component score (RAND-36 PCS). The HADS provided anxiety and depression subscores.

Furthermore, the following demographic data were collected: sex, age, Body Mass Index (BMI), smoking status (yes/no), duration of pain (< 2 years/ ≥ 2 years) and ASA score (I–II/III).

In the validation set, back and leg pain was measured using the 11-point Numeric Pain Rating Scale (NRS) instead of the VAS [9, 14]. The NRS score was transformed to a 0–100 scale by multiplying all scores by ten, to match with the VAS scale in the derivation set.

Dependent variable

Pain relief is the main goal for most patients undergoing LSF [15]. Therefore, the primary outcome of the prediction tool was defined as a clinically relevant reduction in predominant pain in the back or (one of the) legs (worst reported pain in back or legs) as measured with the VAS at 1 to 2 years after surgery. The secondary outcome was defined as a clinically relevant reduction in leg pain at 1 to 2 years after surgery. The VAS for pain ranges from 0 to 100, with 0 indicating no pain and 100 indicating the most severe pain imaginable [9]. To make interpretation of the prediction tool more practical, the dependent variable was made binary: clinically relevant pain reduction or not. Minimal clinically important change (MCIC) for pain ranged between 0.28 and 2.88 on an 11-point scale in the literature on spinal surgery, and a reduction of 2.88 or more (28.8 on a 0–100 point scale) was a priori defined as a clinically relevant pain reduction to prevent overestimation [16].

Statistics

Analyses were performed using SPSS (versions 24, SPSS Inc., Chicago, IL, USA) and R (version 3.3.2; <https://>

www.r-project.org). In the case of incomplete variables within a case, multiple imputation of missing values was used [17].

The independent samples *t* test for normally distributed variables or the Mann–Whitney *U* test for nonnormally distributed variables was used to analyse differences in baseline and outcome variables between subgroups within and between cohorts.

Multivariable logistic regression was used to develop the prediction model. Stepwise backward elimination was used to eliminate nonsignificant predictor variables from the logistic regression model. To prevent premature deletion of predictor variables, a more liberal alpha for exclusion criterion of variables was used ($\alpha = 0.157$) [18].

Discriminatory capacity of the prediction model was quantified by the area under the receiver operating characteristic curve (AUC). The discriminative capacity is perfect when the AUC is 1.0; there is no discriminative capacity when the AUC is 0.5 equivalent to a coin flip.

The logistic regression model was internally validated using standard bootstrapping techniques. As a result, a shrinkage factor was computed, which was used to penalize the regression coefficients of the logistic regression model. The internally validated model was applied to the validation set, for which a new AUC was calculated to evaluate its performance in the population of the second hospital. A nomogram was developed from the validated logistic regression model.

Power analysis

As a general rule, ten events per predictor variable are necessary to find associations in logistic regression models [19]. The percentage of patients undergoing LSF achieving MCIC in pain on average was 56% [3]. A prediction model with 11 predictors could be developed based on a sample size of 197 patients (202 patients were available in the derivation set). Eleven independent variables were selected based on clinical relevance by literature review [4–7] and by expert opinion of five experienced spine surgeons. Selected variables include the following: sex, BMI, pain duration, smoking status, educational level, employment status, ASA score, VAS, ODI, PCS and RAND-36 [4–7].

Results

Population characteristics

The derivation set consisted of 202 patients who were found eligible for analysis (see Fig. 1). Baseline characteristics are shown in Table 1. The mean reduction in predominant pain

Fig. 1 Flowchart of number of patients included in the dataset used to develop the model. *LSF* lumbar spinal fusion

was 33/100 points (SD = 31.3); for leg pain, it was 35/100 (SD = 35.5).

The validation set consisted of 251 patients (see Table 1). The validation set differed from the derivation set in terms of the mean preoperative predominant pain score ($P = 0.001$), RAND-36 MCS ($P = 0.047$) and reduction in predominant pain ($P = 0.044$). The mean reduction in predominant pain in the validation set was 27/100 points (SD = 29.4); for leg pain, this was 31/100 (SD = 34.6). No significant differences in terms of predominant pain reduction were found between categories of surgery type, primary diagnosis or number of levels fused (see Table 2).

Development of the prediction model

In total, 9.1% of values were missing in the derivation set; these values were imputed using 20 imputations.

The clinical prediction model consisted of eight independent predictors after stepwise backward elimination: smoking, BMI, pain duration, educational level, ASA, predominant preoperative pain, physical functioning (ODI), HRQOL related to mental health (RAND-36 MSC). Patients had a higher probability (odds ratio [95% confidence interval]) of achieving a clinically relevant pain reduction if they were nonsmoking patients (0.41 [0.19–0.87]) with lower BMI (0.93 [0.85–1.01]), short pain duration (0.49 [0.20–1.19]), low educational level (0.46 [0.19–1.12]), lower ASA score (4.82 [1.35–17.25]), higher VAS scores (1.05 [1.02–1.08]), lower ODI (0.96 [0.93–1.00]) and higher RAND-36 MCS (1.03 [0.10–1.06]) (see Table 3). The model had an AUC of 0.77 (95% CI = 0.70–0.83).

The model for leg pain consisted of four independent predictors after stepwise backward elimination: smoking,

Table 1 Cohort characteristics and differences between derivation sample, complete case and external validation sample*

Variable	Derivation sample (n = 202)	Complete case sample (n = 73)	External validation sample (n = 251)	Difference complete/incomplete cases derivation sample ^a	Difference derivation/external sample ^a
Age in years, mean (SD)	58 (13.1)	59 (11.4)	57(11,1)	0.519 ^m	0.199 ^m
Sex, % female (n)	59.9% (121)	63.1% (65)	57.8% (145)	0.828 ^m	0.647 ^m
BMI, mean (SD)	27 (4.5)	27 (4.8)	28 (5,0)	0.902 ^t	0.936 ^m
Duration of pain					
< 2 years	40.6% (82)	53.4% (39)	40.6% (102)	0.267 ^m	0.111 ^m
≥ 2 years	43.1% (87)	46.6% (34)	59.4% (149)		
Smoking (% yes)	31.7% (64)	38.4% (28)	35,5% (89)	0.426 ^m	0.054 ^m
Educational level					
Low	38.1% (77)	42.5% (31)	n/a	0.562 ^m	n/a
High	46.5% (94)	57.5% (42)			
Employment status					
Employed	24.8% (50)	27.4% (20)	n/a	0.545 ^m	n/a
Unemployed	62.9% (127)	72.6% (53)			
ASA score, mean (SD)					
I–II	86.6% (175)	87.7% (64)	88.0% (221)	0.745 ^m	0.664 ^m
III	13.4% (27)	12.3% (9)	12.0% (30)		
Pain score (0–100), mean (SD)	74 (17.5)	73 (16.7)	80 (14.1)	0.167 ^m	0.001 ^m
HADS (0–21), mean (SD)					
Anxiety	7 (4.2)	7 (4.2)	n/a	0.790 ^m	n/a
Depression	7 (4.4)	7 (4.5)	n/a	0.491 ^m	
PCS (0–52), mean (SD)	25 (12.3)	24 (12.5)	n/a	0.387 ^m	n/a
ODI (0–100), mean (SD)	45 (12.7)	45 (14.1)	68 (154)	0.580 ^t	0.121 ^t
RAND-36 (0–100), mean (SD)					
PhCS	29 (7.7)	28 (6.1) 46	29 (7.6) 43	0.088 ^t	0.406 ^m
MCS	45 (12.7)	(12.1)	(12.2)	0.559 ^m	0.047 ^m
Change score pain, mean (SD) (–100–100)	–33 (31.3)	–32 (30.9)	–27 (29.4)	0.678 ^m	0.044 ^m
Change score leg pain, mean (SD) (–100–100)	–35 (35.5)	–34 (35.8)	–31 (34.6)	0.701 ^m	0.293 ^m

ASA American Society of Anaesthesiologists; VAS Visual Analogue Scale; HADS Hospital Anxiety and Depression Scale; PCS Pain Catastrophizing Scale; ODI Oswestry Disability Index, RAND-36 PhCS Physical Component Score; RAND-36 MCS Mental Component Score

^aP value: comparison of two samples. Statistical testing of pooled results, independent t test (t), Mann–Whitney U test (m)

*This table provides the mean and standard deviations for continuous variables and the percentage and counts for categorical variables. N=202, derivation sample; for some participants, baseline statistics were missing and are therefore not shown

pain duration, ASA, predominant preoperative pain. Patients had a higher probability of achieving a clinically relevant leg pain reduction if they were nonsmoking (0.55 [0.27–1.12]), had short pain duration (0.59 [0.30–1.15]), lower ASA score (3.18 [0.82–12.34]) and higher VAS scores (1.03 [1.01–1.05]). The model had an AUC of 0.71 (95% CI=0.63–0.77).

Internal validation

The bootstrap validation yielded a shrinkage of 0.84 for predominant pain and 0.88 for leg pain, which was used to multiply the regression coefficients of the final model in

order to correct for overfitting (see Table 4). The optimism-corrected AUC of the internally validated model was 0.74 for predominant pain and 0.69 for leg pain.

External validation

After exclusion of patients who had not completed any preoperative PROMs, 0.18% of the values were missing and these were imputed. Educational level was missing in the validation cohort and was therefore omitted from the prediction model. In the validation set, the prediction model was able to discriminate between achieving relevant pain reduction or not in 68% of the cases, meaning that

Table 2 Surgery characteristics and subgroup distribution with regard to clinically relevant pain reduction 1 to 2 years after LSF

Variable	Derivation sample (<i>N</i> =202), % (<i>N</i>)	External validation sam- ple (<i>N</i> =251), % (<i>N</i>)	Difference derivation/ external validation sample ^a
Type of surgery			0.444 ^k
PLIF ^{b, c}	85.1 (172) 4.0 (8)	100 (251)	
Posterolateral fusion without decompression	10.9 (22)	0 (0)	
TLIF ^{b, c}	0 (0)	0% (0)	
MIS		0% (0)	
Primary diagnosis			0.058 ^k
Deg. with listhesis	73.3 (148)	40.6 (102)	
Deg. without listhesis	14.4 (29)	34.3 (86)	
Prior spine surgery*	4 (8)	12.7 (32)	
Adjacent level disease	8.3 (17)	12.4 (31)	
Levels fused			0.460 ^k
One	81.2 (164)	64.1 (161)	
Two	17.8 (36)	31.1 (78)	
Three	0.5 (1)	4.8 (12)	

PLIF posterior lumbar interbody fusion; TLIF transversal lumbar interbody fusion Deg degenerative disorders of the lumbar spine

^a*P* value: comparison of two samples or more samples. Statistical testing of pooled results Kruskal–Wallis test (*k*)

^bSurgery technique includes usage of interbody cage

^cInherent decompression

*Post-herniotomy/-laminectomy

Table 3 Predictors of a clinically relevant reduction in predominant pain 1 to 2 years after LSF using logistic regression

Variable	Regression coefficient		<i>P</i> value	Odds ratio	95% Confidence interval	
	β	S.E			Lower	Upper
BMI	−0.075	0.045	0.094	0.927	0.848	1.013
Pain duration (long)	−0.718	0.455	0.114	0.487	0.199	1.188
Smoking status (yes)	−0.890	0.386	0.021	0.410	0.192	0.874
Educational level (high)	−0.772	0.454	0.089	0.461	0.189	1.123
ASA score (I–II)	1.574	0.650	0.015	4.825	1.349	17.254
VAS	0.051	0.015	0.000	1.052	1.021	1.083
ODI	−0.035	0.018	0.051	0.965	0.932	1.000
RAND-36 MCS	0.030	0.016	0.067	1.030	0.997	1.063
Constant	−1.558	1.895	0.411	0.210	0.005	8.637

LSF lumbar spinal fusion; ASA American Society of Anaesthesiologists; VAS Visual Analogue Scale; ODI Oswestry Disability Index; RAND-36 MCS 36 Mental Component Scale

an AUC of 0.68 (95% CI=0.66–0.69) was achieved. For leg pain, the AUC in the validation set was 0.52 (95% CI=0.44–0.59).

Development of the prediction tool

From the validated model for clinically relevant reduction in predominant pain, a nomogram was plotted (see Fig. 2). Patients score points per predictor variable, as visualized on the rulers. Explanation on how to use the nomogram and a practical example can be found in “Appendix 1.”

Table 4 Internally validated logistic prediction model for clinically relevant pain reduction 1 to 2 years after LSF

Variable	β -coefficients	
	Predominant pain	Leg pain
BMI	−0.063	n/a
Pain duration (long)	−0.603	−0.468
Smoking status (yes)	−0.748	−0.533
ASA score (I–II)	1.322	1.021
Max. pain	0.043	0.027
ODI	−0.029	n/a
RAND-36 MCS	0.025	n/a
Constant	−2.466	−3.679

LSF lumbar spinal fusion; ASA American Society of Anaesthesiologists; VAS Visual Analogue Scale; ODI Oswestry Disability Index; RAND-36 MCS 36 Mental Component Scale

Sensitivity analysis

Primary diagnosis, as categorized in Table 2, was added as a predictor to the clinical prediction model, to assess whether variability in diagnosis within our population influenced the final prediction model. Primary diagnosis was excluded from the final prediction model after stepwise backward elimination.

Discussion

We developed and validated a tool to preoperatively predict a clinically relevant reduction 1 to 2 years after LSF in an adequately powered analysis. A nomogram was developed from the externally validated model (for the primary outcome) for application in clinical practice. With an AUC of 0.68 in an external population, this prediction tool possesses fair discriminatory ability to predict a clinically relevant reduction in predominant pain. We also developed and externally validated a model for clinically relevant reduction in leg pain, which had an AUC of 0.52 and thus possesses low discriminatory ability. The clinical prediction tool for predominant pain could be implemented in clinical practice to improve shared decision making when considering LSF.

In agreement with our findings, previous studies reported that preoperative nonsmoking status [5, 7], better physical functioning [4, 6] and better mental health [4, 5] predict pain reduction 1 to 2 years after LSF. This strengthens the likelihood that the prediction tool developed in this study is able to predict pain reduction in other populations as well.

Surprisingly, our results showed that higher educational level indicated a lower probability of a clinically relevant pain reduction, whereas from the literature high socioeconomic status is usually associated with a better health condition, especially in patients with chronic low back pain [20, 21]. Educational systems in various countries are different,

Fig. 2 Nomogram to predict the probability of a clinically relevant pain reduction for the individual patient. BMI Body Mass Index; ASA American Society of Anaesthesiologists; ODI Oswestry Disability Index; MCS Mental Component Scale

and definitions of high educational level can differ; therefore, further research is needed to verify this finding.

The performance of our prediction nonvalidated model for reduction in predominant pain was similar to that of Abbott et al. (0.74 vs. 0.72 respectively); the externally validated model of Kohr et al. performed better compared to ours (0.79 vs. 0.68 respectively) [7, 22]. However, they externally validated their model in a random sample from the same population it was built in, explaining the high performance. The model performance for reduction in leg pain was low (AUC=0.52). Therefore, this model was not translated into a prediction tool. A possible explanation for the low AUC is that we excluded possibly important predictors too soon in the model development phase, leading to overfitting (data fits "too well") of the model to the derivation set [18]. The added value of our study lies in the fact that we externally validated a model predicting a clinically relevant reduction in predominant pain in a European setting and translated it into a concrete tool for use in clinical practice (see Fig. 2).

Strengths and limitations

A strength of the study is that our model is derived from an academic hospital population and externally validated on a population from a general hospital. Usually, surgical populations from an academic hospital and general hospital differ in the sense of complexity of the surgery. From our external validation, it is apparent the model can predict a clinically relevant reduction in predominant pain in both academic (AUC=0.74) and nonacademic settings (AUC=0.68). However, for leg pain, this was not the case as it did not perform well in the nonacademic setting (AUC=0.52). Further external validation of the prediction tool is necessary for applicability of the prediction tool to countries with different surgical populations and healthcare systems.

A limitation of this study is the amount of missing data in derivation set used to develop the model (9.1%). This was probably caused by the fact that the data were collected retrospectively from standard care records. Consequently, multiple imputation was to minimize to increase the power of our analysis. Secondly, in the general hospital, the variable "educational level" was missing [23]. We chose elimination of this predictor from the model rather than imputation, because the value of this predictor is considered untrustworthy without external validation. Finally, we acknowledge that the cutoff point for clinical relevance in our model, although based on literature, is arbitrary. Nevertheless, the primary outcome was defined as a clinically relevant reduction in predominant pain, as indications for elective LSF are due to both back and leg pain in our hospitals.

Future implications of the results

The validated prediction tool for estimating clinically relevant reduction in predominant pain can be used by clinicians as an aid to preoperatively inform individual patients about their expected outcomes. An example and explanation of the clinical application and decision making with the help of nomogram can be found in "Appendix 1." Secondly, adding new variables able to predict clinically relevant pain reduction could improve the performance of the prediction models. A variable that is overlooked in all previously mentioned models is preoperative physical performance. In other types of surgery, it has been proven physical performance can improve predictive power [24, 25], which may also hold true for patients undergoing LSF. Thirdly, for patients who are less likely to achieve a clinically relevant pain reduction, care should be tailored to their specific needs in order to improve this probability. Using the nomogram, a surgeon can identify which risk factors that are modifiable contribute least to the expected pain reduction for the individual patient and can inform the patient to improve these risk factors before surgery.

Conclusion

Using the validated prediction tool (nomogram), a patient's probability of a clinically relevant pain reduction can be estimated 1 to 2 years after undergoing LSF. This validated prediction tool can be implemented in clinical practice to aid patients and care professionals in the difficult process of clinical decision making when considering LSF.

Funding No funding.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References:

1. Yoshihara H, Yoneoka D (2015) National trends in the surgical treatment for lumbar degenerative disc disease: United States, 2000 to 2009. *Spine J* 15(2):265–271
2. Bogduk N, Andersson G (2009) Is spinal surgery effective for back pain? *F1000 Med Rep* 1:S2–S3
3. Glassman SD, Carreon LY, Djurasovic M, Dimar JR, Johnson JR, Puno RM, Campbell MJ (2009) Lumbar fusion outcomes stratified by specific diagnostic indication. *Spine J* 9(1):13–21
4. Abbott AD, Tyni-Lenne R, Hedlund R (2011) Leg pain and psychological variables predict outcome 2–3 years after lumbar fusion surgery. *Eur Spine J* 20(10):1626–1634
5. Trief PM, Ploutz-Snyder R, Fredrickson BE (2006) Emotional health predicts pain and function after fusion: a prospective multicenter study. *Spine* 31(7):823–830
6. Ekman P, Moller H, Hedlund R (2009) Predictive factors for the outcome of fusion in adult isthmic spondylolisthesis. *Spine* 34(11):1204–1210
7. Khor S, Lavalley D, Cizik AM, Bellabarba C, Chapman JR, Howe CR, Lu D, Mohit AA, Oskouian RJ, Roh JR, Shonnard N, Dagal A, Flum DR (2018) Development and validation of a prediction model for pain and functional outcomes after lumbar spine surgery. *JAMA Surg* 153(7):634–642
8. Toll DB, Janssen KJ, Vergouwe Y, Moons KG (2008) Validation, updating and impact of clinical prediction rules: a review. *J Clin Epidemiol* 61(11):1085–1094
9. Downie WW, Leatham PA, Rhind VM, Wright V, Branco JA, Anderson JA (1978) Studies with pain rating scales. *Ann Rheum Dis* 37(4):378–381
10. Fairbank JC, Couper J, Davies JB, O'Brien JP (1980) The Oswestry low back pain disability questionnaire. *Physiotherapy* 66(8):271–273
11. Hays RD, Morales LS (2001) The RAND-36 measure of health-related quality of life. *Ann Med* 33(5):350–357
12. Sullivan MJ, Bishop SR, Pivik J (1995) The pain catastrophizing scale: development and validation. *Psychol Assess* 7(4):524
13. Zigmond AS, Snaith RP (1983) The hospital anxiety and depression scale. *Acta Psychiatr Scand* 67(6):361–370
14. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole RM (2001) Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* 94(2):149–158
15. Mancuso CA, Reid MC, Duculan R, Girardi FP (2017) Improvement in Pain after lumbar spine surgery. *Clin J Pain* 33(2):93
16. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY (2008) Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry disability index, medical outcomes study questionnaire short form 36, and pain scales. *Spine J* 8(6):968–974
17. Held U, Kessels A, Garcia Aymerich J, Basagana X, Ter Riet G, Moons KG, Puhan MA (2016) Methods for handling missing variables in risk prediction models. *Am J Epidemiol* 184(7):545–551
18. Steyerberg E (2008) Clinical prediction models: a practical approach to development, validation, and updating. Springer Science & Business Media, Berlin
19. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR (1996) A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 49(12):1373–1379
20. Katz JN (2006) Lumbar disc disorders and low-back pain: socioeconomic factors and consequences. *JBJS* 88:21–24
21. Ross CE, Wu C-I (1995) The links between education and health. *Am Sociol Rev* 60:719–745
22. Abbott AD, Tyni-Lenne R, Hedlund R (2011) Leg pain and psychological variables predict outcome 2–3 years after lumbar fusion surgery. *Eur Spine J* 20(10):1626–1634
23. Bennett DA (2001) How can I deal with missing data in my study? *Aust N Z J Public Health* 25(5):464–469
24. Punt IM, Bongers BC, Van Beijsterveld C, Hulzebos H, Dronkers J, Van Meeteren N (2016) Surgery: moving people, improving outcomes. *Geriatr Hyderabad: Avid Sci* 1:2–29
25. Hulzebos E, van Meeteren N (2016) Making the elderly fit for surgery. *Br J Surg* 103(2):e12–e15

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.