ELSEVIER

The Spine Journal 000 (2019) 1–8

THE SPINE JOURNAL

Clinical Study

Predictors of loss of follow-up in a prospective registry: which patients drop out 12 months after lumbar spine surgery?

Marc L. Schröder, MD, PhD[a], Marlies P. de Wispelaere, MSc, PDEng[b], Victor E. Staartjes, BMed[a,c,*]

[a] *Department of Neurosurgery, Bergman Clinics Amsterdam, Rijksweg 69, 1411 GE, Naarden, The Netherlands*
[b] *Department of Clinical Informatics, Bergman Clinics, Gooimeer 11, 1411 GE, Naarden, The Netherlands*
[c] *Department of Neurosurgery, Clinical Neuroscience Center, University Hospital Zurich, University of Zurich, Frauenklinikstrasse 10, CH-8091, Zurich, Switzerland*

Received 19 February 2019; revised 13 May 2019; accepted 14 May 2019

Abstract

BACKGROUND CONTEXT: Long-term patient-reported outcomes (PROMs) are essential in clinical practice and research. Prospective trials and registries often struggle with high rates of loss of follow-up (LOFU), which may bias their findings. Little is known on risk factors for PROM non-response, especially for digitally mailed questionnaires.
PURPOSE: To elucidate which patients are at high risk for LOFU by identifying associated predictors.
STUDY DESIGN: Analysis of a prospective registry.
PATIENT SAMPLE: Patients that underwent surgery for degenerative lumbar disease were included.
OUTCOME MEASURES: Rate of PROM follow-up response at 12 months postoperatively.
METHODS: Preoperatively and at 12 months postoperatively, patients were asked to complete a range of PROM questionnaires using a web-based tool. All patients who successfully completed their baseline questionnaire were included. Patients were not actively reminded upon nonresponse. Univariate and independent predictors of LOFU at 12 months were identified.
RESULTS: We included 1,456 patients, of which 861 (59%) were lost to follow-up at 12 months. Univariately, lower age, American Society of Anesthesiologists (ASA) class 1, smoking, lack of prior surgery, higher pain scores and functional disability, and lower quality-of-life were associated with LOFU (all $p<.05$). Only lower age (OR: 0.98, p=.001), smoking (OR: 1.46, p=.019), lack of prior surgery (OR: 0.59, p=.019), and spondylolisthesis (OR: 0.47, p=.024) independently predicted LOFU.
CONCLUSIONS: In a prospective registry of lumbar spine surgery patients based on web-based outcome capturing, younger age, active smoking status, lack of prior surgery, and nonspondylolisthesis surgery were independent predictors of loss of follow-up. In the future, it may become possible to preoperatively identify patients at high-risk for study dropout. As the implementation of prospective registries and the use of automated follow-up methods are on the rise, it is crucial to ensure efficiency and reduce bias of the methods on which all clinical research is based on.

Keywords: Outcome measurement; Registry; Loss of follow-up; Dropout; Patient-reported outcome measures; Risk factors

This research has never previously been submitted for review or presented at any conferences.

FDA device/drug status: Not applicable.

Author disclosures: ***MS:*** Nothing to disclose; ***MW:*** Nothing to disclose; ***VS:*** Nothing to disclose.

Level of evidence: 2

Ethical approval: This prospective registry was approved by the local institutional review board (Medical Research Ethics Committees United, Registration Number: W16.065), and was conducted according to the Declaration of Helsinki. Informed consent was obtained from all participants.

* Corresponding author. Bergman Clinics, Naarden, 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.2019.05.007
1529-9430/

Introduction

Patient-reported outcome measures (PROMs) are essential in research and clinical follow-up in day-to-day spinal patient care. Consequently, a large number of institutional, national, and international registries have been initiated to assess the effectiveness of treatments using PROMs as a measure of success [1]. In contrast to clinical trials, registries are able to capture real-world care in real-world clinical situations with large sample sizes allowing for powerful analysis. Other than for research, it has been reported that around 70% of spine surgeons worldwide regularly use PROMs in their clinical practice [2].

There is a consensus that one- or two-year PROMs constitute the minimum follow-up length in studies dealing with degenerative lumbar spine surgery [3–6]. Short-term PROM's may not accurately reflect real-world clinical value of surgical treatments in the longer term [4,7]. However, it is logistically and financially not always feasible to capture long-term PROM's. Especially at long-term follow-up, the results of prospective studies are often hampered by up to 78% of loss of follow-up (LOFU), depending among other factors on the method of PROM capturing [1,6,8].

Little is known on factors affecting the rate of long-term follow-up after spine surgery [9–13]. In prospective clinical studies, where every recruited patient counts, independent predictors of LOFU could aid in identifying those patients that are at high risk of dropping out, and may thus require tighter monitoring to ensure complete follow-up. Additionally, dropout of specific subsets of patients may constitute a significant bias in prospective clinical studies. Specifically, studies on registries that use web-based PROM collection, and that target independent predictors of LOFU have been called for [13]. For these reasons, the purpose of this study was to elucidate which patients are at high risk for LOFU by identifying associated predictors in a prospective registry.

Materials and methods

Study design

To identify predictors of LOFU after lumbar spine surgery, we analyzed patients in a prospective registry. Those patients who returned their 12-month follow-up questionnaires were contrasted with those who did not. The study was compiled according to the Strengthening the Reporting of Observational Studies in Epidemiology statement [14]. This registry was approved by the local institutional review board (Medical Research Ethics Committees United, Registration Number: W16.065), and this study was conducted according to the Declaration of Helsinki. Informed consent was obtained from all individual participants.

Patient population

Exclusion criteria were pediatric and trauma patients and those older than 80 years, as well as red flags such as malignancies or severe comorbidities. In addition, we did not include patients with less than 80% baseline data completion. Indications for surgery were lumbar disc herniation (LDH), lumbar spinal stenosis (LSS), degenerative or isthmic spondylolisthesis, discopathy, and synovial facet cysts [15–20]. Discopathy was defined as chronic low back pain caused by single- or double-level degenerative disc disease [21]. All patients were operated at a specialized spine surgery center using minimally-invasive or mini-open techniques. Patients underwent either tubular microdiscectomy, mini-open laminectomy, minimally-invasive robot-guided interbody fusion, mini-open anterior lumbar interbody fusion, or minimally-invasive transaxial interbody fusion [21–23]. Patients scheduled for fusion surgery were required to cease smoking. Complications and reoperations were systematically tracked.

Outcome measures

Baseline and follow-up PROM were obtained using web-based questionnaires [24]. These questionnaires consisted of an e-mailed invitation to complete a web-based PROM survey on software designed specifically for this purpose by our Department of Clinical Informatics [24]. Patients were not actively reminded to fill out the questionnaires upon non-completion. The questionnaires included Numeric Rating Scales (NRS) for back and leg pain severity, and validated Dutch versions of the Oswestry Disability Index for functional disability as well as EuroQOL-5D-3L index (EQ-5D index) and visual analogue scale (EQ-VAS) for health-related quality of life (HRQOL) [25,26]. The EQ-5D was evaluated according to the Dutch tariff [25]. As a safety criterion, we only included patients who had filled out their basic mailed preoperative PROM questionnaire containing NRS for back and leg pain as well as Oswestry Disability Index. This allowed us to ensure that all included patients were reached and were able to complete the questionnaires. Patients received the same mailed questionnaires at 6 weeks, 12 months, and 24 months. Consequently, we defined the primary endpoint of this study as LOFU at 12 months postoperatively.

Statistical analysis

Continuous data are reported as mean ± standard deviation, and categorical data as numbers (percentages). Intergroup differences in baseline characteristics were assessed among subgroups and among the included vs. excluded patients by use of Welch's two-sample *t* or chi-square tests. A multivariate logistic regression model was fitted to identify factors independently associated with LOFU, based on the Hosmer–Lemeshow "purposeful selection of variables" procedure, as described by Bursac et al. [27]. In more detail, variables were considered for inclusion at univariate $p \leq .25$. Subsequently, a multivariate model was built, and variables that did not have a significant effect (defined as $p \leq .1$) or that did not demonstrate confounding (defined

using the change-in-estimate criterion of 20% or more) were iteratively removed from the model. Finally, any variable not eligible for the original multivariate model was added iteratively, and the model was subsequently reduced in the same way as described above by iterative removal of only those variables that were additionally added [27]. We conducted a posthoc power analysis based on the results of the multivariate regression model. Missing data were encountered for the following variables: smoking status, American Society of Anesthesiologists (ASA) score, body mass index, and HRQOL. These variables were assumed to be missing at random. We used multiple imputation by chained equations to impute missing baseline data. Predictive mean matching was used as the imputation method [28,29]. A Benjamini–Hochberg correction for multiple testing was consistently applied to control the false discovery rate while retaining power [30]. A sensitivity analysis (Supplementary Content 1), based on modeling with only those variables with univariate statistical significance, is also provided. All analyses were carried out using version 3.5.1 of R (The R Foundation for Statistical Computing, Vienna Austria) [31]. A two-tailed $p \leq .05$ was considered significant.

Results

Patients

The flow of patients throughout this analysis is reported in Fig. 1. Between May 2013 and April 2018, 1,456 patients (100%) responded to the initial baseline questionnaire. Detailed patient characteristics are given in Table 1. Most patients presented with LDH (1097 pts., 75%) or LSS (229 pts., 16%). Multilevel surgery was only seen in patients undergoing decompression for LSS (77 pts., 6%). Baseline data were 89.5% complete. Of the 1,456 included patients who had completed and returned their baseline questionnaire, 595 (41%) completed and returned the mailed 12-month PROM questionnaire. Accordingly, the rate of LOFU at 12 months was 59% (861 pts.). The 6-week and 24-month PROM follow-up rates were 62% (900 pts.) and 21% (312 pts.), respectively. Our posthoc power analysis indicated that the power of our primary analysis approached one.

Univariate predictors of loss of follow-up

Variables were contrasted between patients who did and who did not return the 12-month questionnaire (Table 2). Lower age (Δ: −4.2 years, 95% CI: −5.5 to −2.9, p<.001), ASA class 1 (61% vs. 54%, p=.019), active smoking status (33% vs. 26%, p=.009), lack of prior surgery (8% vs. 12%, p=.02), higher NRS leg pain (Δ: 0.4, 95% CI: 0.1–0.6, p=.019) and back pain (Δ: 0.4, 95% CI: 0.1–0.7, p=.02) severity, higher functional disability (Δ: 2.8, 95% CI: 0.9–4.7, p=.019), and lower EQ-5D index HRQOL status (Δ: −0.07, 95% CI: −0.1 to −0.03, p=.047) were associated

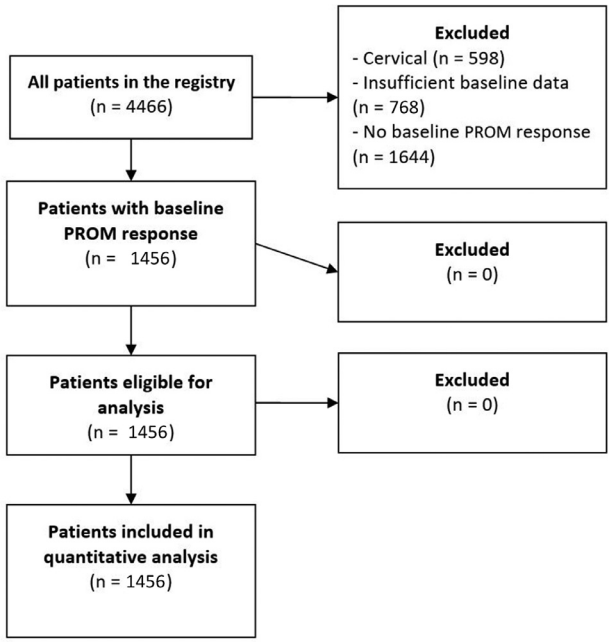

Fig. 1. Flow chart illustrating the flow of patients throughout this study.

with LOFU. In addition, the indication for surgery affected the rate of LOFU (p=.008), with LDH (62%) and synovial facet cysts (67%) associated with greater LOFU than spondylolisthesis (39%), discopathy (50%), and LSS (53%). When

Table 1
Baseline characteristics of the patient cohort

Parameter	Value N=1,456
Male gender, n (%)	788 (54)
Age, mean±SD [yrs.]	48.5±13.0
BMI, mean±SD [kg/m^2]	25.6±3.3
ASA class I, n (%)	819 (58)
Active smoker, n (%)	438 (30)
Prior surgery, n (%)	142 (10)
Indication, n (%)	
Disc herniation	1,097 (75)
Stenosis	229 (16)
Single-level decompression	152 (10)
Multi-level decompression	77 (6)
Spondylolisthesis	56 (4)
Discopathy	62 (4)
Synovial facet cyst	12 (1)
Primary Index level	
L1-L2	4 (0)
L2-L3	43 (3)
L3-L4	167 (12)
L4-L5	621 (43)
L5-S1	621 (43)
Baseline PROMs, mean±SD	
NRS leg pain severity	7.1±2.3
NRS back pain severity	5.6±2.8
Oswestry Disability Index	46.9±18.0
EQ-5D index	0.40±0.31
EQ-VAS	50.2±18.6

SD, standard deviation; BMI, body mass index; ASA, American Society of Anesthesiologists; NRS, numeric rating scale.
The analysis was performed on the multiply imputed dataset.

Table 2
Univariate analysis of factors associated with loss of follow-up

Parameter	Successful FU N=595	Loss of FU N=861	p
Male gender, n (%)	318 (53)	470 (55)	.711
Age, mean±SD [y]	51.0±12.3	46.8±13.2	<.001*
BMI, mean±SD [kg/m^2]	25.8±3.3	25.4±3.3	.121
ASA class I, n (%)	322 (54)	527 (61)	.019*
Active smoker, n (%)	151 (26)	285 (33)	.009*
Prior surgery, n (%)	70 (12)	70 (8)	.02*
Indication, n (%)			
Disc herniation	419 (70)	678 (79)	.008*
Stenosis	108 (18)	121 (14)	
Spondylolisthesis	34 (6)	22 (3)	
Discopathy	31 (5)	31 (4)	
Synovial facet cyst	4 (1)	8 (1)	
Primary Index level			
L1-L2	2 (0)	2 (0)	.235
L2-L3	22 (4)	21 (2)	
L3-L4	69 (12)	98 (11)	
L4-L5	269 (45)	352 (41)	
L5-S1	233 (39)	388 (45)	
Baseline PROMs, mean±SD			
NRS leg pain severity	6.9±2.4	7.3±2.2	.019*
NRS back pain severity	5.3±2.9	5.7±2.8	.02*
Oswestry Disability Index	45.2±18.6	48.0±17.6	.019*
EQ-5D index	0.44±0.32	0.37±0.30	.047*
EQ-VAS	51.0±19.0	49.7±18.4	.235

FU, follow-up; SD, standard deviation; BMI, body mass index; ASA, American Society of Anesthesiologists; NRS, Numeric Rating Scale.

* p≤.05.

The analysis was performed on the multiply imputed dataset.

statistically comparing these univariate predictors among patients included against those excluded in the analysis caused by initial nonresponse (n=1,644) to the baseline PROM questionnaire, we did not identify differences in age (p=.410), active smoking status (p=.251), prior surgery (p=.637), and indication for surgery (p=.283). However, the excluded population exhibited a reduced ASA class 1 ratio (58% vs. 63%, p=.007).

Independent predictors of loss of follow-up

A logistic regression model, which included all significant variables from the univariate analysis, was trained and evaluated on the entire imputed dataset (Table 3). We identified lower age (odds ratio (OR): 0.98, 95%: 0.97−0.99, p<.001), active smoking status (OR: 1.43, 95% CI: 1.11−1.84, p=.016), lack of prior surgery (OR: 0.54, 95% CI: 0.37−0.78, p=.007), as independent predictors of LOFU. Additionally, patients who presented with spondylolisthesis were independently at lower risk for LOFU (OR: 0.48, 95% CI: 0.27−0.84, p=.029), compared with LDH as the reference category. Fig. 2 demonstrates this and illustrates that fusion procedures (discopathy and spondylolisthesis) had markedly lower rates of LOFU than procedures without fusion. The multivariate model achieved an area under the receiver operating characteristics curve of 0.635 and Nagelkerke's pseudo-R^2 of 0.078, indicating only moderate predictive ability for LOFU. The sensitivity analysis confirmed these findings.

Discussion

In 1,456 patients from a prospective registry, we identified predictors of LOFU. Younger age, absence of comorbidities, active smoking status, lack of prior surgery, and higher baseline pain and disability scores were univariately associated with dropout. In a multivariate analysis, we identified, lower age, active smoking status, lack of prior surgery, as well as absence of spondylolisthesis as the indication for surgery as independent predictors of LOFU.

High rates of LOFU are common in large prospective trials and registries. In the literature, rates ranging from 21% to 78%, with longer follow-up times consistently leading to higher rates of LOFU [1]. In some cases, where LOFU is limited, modern implementations of multiple imputation can be applied to replace missing baseline or endpoint data [29]. Here, the rationale for imputation is primarily to preserve statistical power, but also to prevent biases created by missing demographic subsets of patients. In fact, it has been demonstrated that imputation, in comparison to complete case analysis, is beneficial in clinical spinal research [28]. Nonetheless, imputing large amounts of data in registries with high rates of LOFU is still considered unreliable. In these cases, it remains important to either minimize dropout by more tightly monitoring patients at high risk of LOFU, or to consider adjusting for the dropout of specific patient demographics with increased LOFU, which may otherwise bias findings [10]. Both strategies require information on reliable predictors of LOFU.

In our registry, younger patients were more likely to not return their questionnaires, which may be explained by better preoperative heath status, and less comorbidities. In comparison to their older counterparts, these factors appear to be the main drivers of nonresponse at long-term follow-up. It is also conceivable that some bias is created by the use of digital questionnaires as opposed to paper-based, mailed follow-up, as this assumes a level of digital proficiency as well as motivation. Although it has been demonstrated that reported outcomes do not differ among these two types of follow-up assessment, the effect size of this potential bias on response rates remains unclear in the spine literature [24]. Similarly, patients who underwent surgery for disc herniation or facet joint cysts in our population were less likely to respond at 1 year postoperatively compared with those who underwent fusion for spondylolisthesis or clinical mono-discopathy. Again, the differential success rates, recovery speed, and length of hospital stay among the surgical indications may explain this phenomenon, although its causality has not been explicitly studied [32−34]. In addition, it is not inconceivable that LOFU may be related to the amount of postoperative follow-up for each type of procedure.

Active smokers were also more likely to drop out. Although it is not inconceivable that this effect is partially

Table 3
Results of the multivariate logistic regression analysis based on purposeful selection of variables as described by Hosmer and Lemeshow. Parameters independently associated with loss of follow-up were identified. The most common indication, namely disc herniation, was used as the reference level to maximize power

Parameter N=1,456	Odds ratio	95% CI	p	AUC	Pseudo-R^2 (Nagelkerke)
Age	0.98	0.97–0.99	<.001*	0.635	0.078
ASA class 1	1.16	0.92 – 1.47	.345		
Active smoker	1.43	1.11–1.84	.016*		
Prior surgery	0.54	0.37–0.78	.007*		
NRS leg pain severity	1.03	0.97–1.08	.367		
NRS back pain severity	1.05	1.00–1.09	.056		
Oswestry Disability Index	1.00	0.99–1.01	.367		
EQ-5D Index	0.75	0.45–1.24	.367		
EQ-VAS	1.01	1.00–1.02	.113		
Indication					
Disc herniation	Reference	–			
Stenosis	1.13	0.81–1.58	.486		
Spondylolisthesis	0.48	0.27–0.84	.029*		
Discopathy	0.61	0.36–1.06	.147		
Synovial facet cyst	1.84	0.55–7.14	.367		

ASA, American Society of Anesthesiologists; NRS, numeric rating scale; AUC, area und the receiver operating characteristics curve.
* p≤.05.
The analysis was performed on the multiply imputed dataset.

produced by habits associated with smoking, the main explanation is more likely that smokers are at higher risk of comorbidities, and usually present with a worse preoperative disease state [35]. Lastly, prior surgery led to lower rates of LOFU. Patients with persisting symptoms after surgery which eventually require reoperation may be more focused on monitoring their disease state in the long-term than those who experienced an immediate resolution of symptoms after the initial surgery.

Little is known on reasons for and predictors of LOFU. Male gender, active smoking status, younger age, living alone, low socioeconomic status, poor preoperative health, more serious injury, geographic relocation, and dissatisfaction with treatment are commonly associated with LOFU

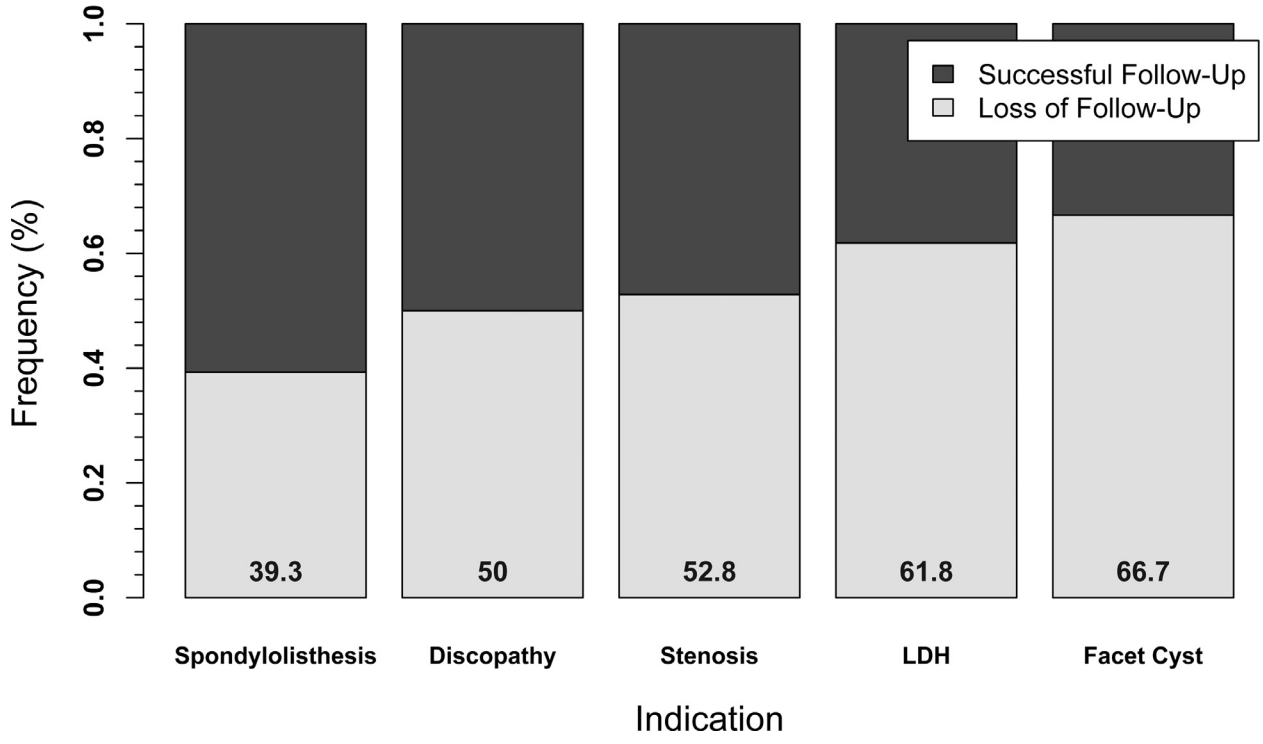

Fig. 2. Bar chart demonstrating the rates of loss of follow-up at 12 months for the various surgical indications. LDH, lumbar disc herniation.

[8–10,12,36]. In spinal patient care, Sielatycki et al. [13] found that smoking was univariately associated with LOFU. The 16% of patients who dropped out were younger and more likely to be preoperatively employed. Their study differs from ours in that they actively followed-up patients using telephone calls or obtained PROM measurements during clinical visits, whereas we employed digitally dispatched questionnaires. In addition, they performed their analysis for cervical and lumbar procedures together and had a relevantly lower rate of 12-month LOFU than in our study. Additionally, the hypothesis that internet-based mailed questionnaires may lead to reduced LOFU in the young patient population is not supported by our findings [13]. In particular, our study determines predictors of 12-month dropout for those patients who have demonstrably agreed to participate in the PROM collection process by completing the baseline PROM assessment, instead of for all patients. First, this leads to restriction to the domain, an epidemiological technique that enables reduced confounding [37]. Second, this more closely represents the situation observed in prospective clinical studies or prospective registries on PROMs: many patients are screened, of which only a subset are included or agree to participate. Subsequently, some of the participants drop out of the study or registry [38]. Even when considering these methodological differences, the same risk factors are consistent across multiple independent studies [8–10,12,13,36], which is a strong token for their effect size and robustness.

Such independent risk factors may be useful in clinical trials, where LOFU of any patient is critical and costly. A patient sporting multiple of these risk factors may benefit from more precise instructions and tighter monitoring, for example using scheduled telephone calls, to minimize the risk of dropout. With further research and machine learning algorithms, it may even become feasible to predict the individual risk of LOFU [39]. However, such promising methodological developments must not go as far as to exclude patients with a high predicted risk of LOFU from enrolling in clinical trials. Aside from the obvious ethical aspect, this would also introduce selection bias through exclusion of certain patient demographics which may then be underrepresented.

As we demonstrate, selection bias is inevitably present because LOFU is itself influenced by baseline characteristics. There is some evidence that dropout influences study outcomes, and it has been historically assumed that patients who drop out are more likely to have experienced a negative outcome [13,40]. However, recent data have challenged this notion. Solberg et al. report that there was no difference in outcome between responders and nonresponders after degenerative spine surgery [36]. Højmark et al. came to a similar conclusion [8]. In fact, patients who are lost to follow-up are often more satisfied and have to undergo less revision surgeries, as Joshi et al. show in knee arthroplasty [11]. However, it is to be noted that both studies had relatively low rates of LOFU with 22% and 12%, respectively. A higher proportion of missing data, as is often seen in large studies, may increase the impact of LOFU on study outcome [1]. If further analyzed in other cohorts, reliable baseline predictors of LOFU may even have the potential to be used for balancing and statistical adjustment in clinical studies with high rates of dropout [10].

Limitations

The registry contained comparatively healthy individuals, without severe comorbidities. In addition, no pediatric, oncological, and trauma patients were included. This means that our conclusions may not be generalizable to those patient populations. We were unable to report on reasons for long-term LOFU, because these data were not regularly captured. Lastly, our PROM data was obtained via mailed digital questionnaires at all timepoints, and not via clinically administered or mailed paper questionnaires. It is arguable whether the method of follow-up influences reported outcomes [24]. Nonetheless, predictors of LOFU may differ among patients who are asked to complete a questionnaire in a clinical setting, vs. those who receive questionnaires at home. However, by only including those patients that successfully completed their mailed baseline questionnaire, it was ensured that all patients could be reached at follow-up and that they were able to complete the mailed questionnaires correctly. However, this meant that exclusion of a large amount of patients may bias our results. To quantify these potential biases, we statistically compared the characteristics of included vs. excluded patients and identified no differences with the exception of a minor discrepancy in ASA scores among excluded patients. We did not have data available on employment status, education levels, income, or job satisfaction, all of which may relevantly influence LOFU.

Conclusions

In a prospective registry of lumbar spine surgery patients based on web-based outcome capturing, younger age, active smoking status, lack of prior surgery, and nonspondylolisthesis surgery were independent predictors of loss of follow-up at 1 year postoperatively. In the future, it may become possible to preoperatively identify patients at high-risk for study dropout. As the implementation of prospective registries, the number of patients per study, and the use of automated follow-up methods are on the rise, it is crucial to ensure efficiency and reduce bias of the methods on which all clinical research is based on.

Supplementary materials

Supplementary material associated with this article can be found in the online version at https://doi.org/10.1016/j.spinee.2019.05.007.

References

[1] McGirt MJ, Parker SL, Asher AL, Norvell D, Sherry N, Devin CJ. Role of prospective registries in defining the value and effectiveness of spine care. Spine 2014;39:S117. https://doi.org/10.1097/BRS.0000000000000552.

[2] Falavigna A, Dozza DC, Teles AR, Wong CC, Barbagallo G, Brodke D, et al. Current status of worldwide use of patient-reported outcome measures (PROMs) in spine care. World Neurosurg 2017;108:328–35. https://doi.org/10.1016/j.wneu.2017.09.002.

[3] Adogwa O, Elsamadicy AA, Han JL, Cheng J, Karikari I, Bagley CA. Do measures of surgical effectiveness at 1 year after lumbar spine surgery accurately predict 2-year outcomes? J Neurosurg Spine 2016;25:689–96. https://doi.org/10.3171/2015.8.SPINE15476.

[4] Asher AL, Chotai S, Devin CJ, Speroff T, Harrell FE, Nian H, et al. Inadequacy of 3-month Oswestry Disability Index outcome for assessing individual longer-term patient experience after lumbar spine surgery. J Neurosurg Spine 2016;25:170–80. https://doi.org/10.3171/2015.11.SPINE15872.

[5] Glassman SD, Schwab F, Bridwell KH, Shaffrey C, Horton W, Hu S. Do 1-year outcomes predict 2-year outcomes for adult deformity surgery? Spine J Off J North Am Spine Soc 2009;9:317–22. https://doi.org/10.1016/j.spinee.2008.06.450.

[6] Staartjes VE, Siccoli A, de Wispelaere MP, Schröder ML. Patient-reported outcomes unbiased by length of follow-up after lumbar degenerative spine surgery: do we need 2 years of follow-up? Spine J Off J North Am Spine Soc 2018. https://doi.org/10.1016/j.spinee.2018.10.004.

[7] Parker SL, Asher AL, Godil SS, Devin CJ, McGirt MJ. Patient-reported outcomes 3 months after spine surgery: is it an accurate predictor of 12-month outcome in real-world registry platforms? Neurosurg Focus 2015;39:E17. https://doi.org/10.3171/2015.9.FOCUS15356.

[8] Højmark K, Støttrup C, Carreon L, Andersen MO. Patient-reported outcome measures unbiased by loss of follow-up. Single-center study based on DaneSpine, the Danish spine surgery registry. Eur Spine J 2016;25:282–6. https://doi.org/10.1007/s00586-015-4127-3.

[9] Daffner SD, Hilibrand AS, Riew KD. Why are spine surgery patients lost to follow-up? Glob Spine J 2013;3:15–20. https://doi.org/10.1055/s-0033-1337120.

[10] Hutchings A, Neuburger J, Grosse Frie K, Black N, van der Meulen J. Factors associated with non-response in routine use of patient reported outcome measures after elective surgery in England. Health Qual Life Outcomes 2012;10:34. https://doi.org/10.1186/1477-7525-10-34.

[11] Joshi AB, Gill GS, Smith PL. Outcome in patients lost to follow-up. J Arthroplasty 2003;18:149–53. https://doi.org/10.1054/arth.2003.50061.

[12] Kim H, Cutter GR, George B, Chen Y. Understanding and preventing loss to follow-up: experiences from the spinal cord injury model systems. Top Spinal Cord Inj Rehabil 2018;24:97–109. https://doi.org/10.1310/sci2402-97.

[13] Sielatycki JA, Parker SL, Godil SS, McGirt MJ, Devin CJ. Do patient demographics and patient-reported outcomes predict 12-month loss to follow-up after spine surgery? Spine 2015;40:1934–40. https://doi.org/10.1097/BRS.0000000000001101.

[14] Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, et al. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007;335:806–8. https://doi.org/10.1136/bmj.39335.541782.AD.

[15] Weinstein JN, Tosteson TD, Lurie JD, Tosteson ANA, Hanscom B, Skinner JS, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the spine patient outcomes research trial (SPORT): a randomized trial. JAMA 2006;296:2441–50. https://doi.org/10.1001/jama.296.20.2441.

[16] Weinstein JN, Lurie JD, Tosteson TD, Skinner JS, Hanscom B, Tosteson ANA, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. JAMA 2006;296:2451–9. https://doi.org/10.1001/jama.296.20.2451.

[17] Ghogawala Z, Dziura J, Butler WE, Dai F, Terrin N, Magge SN, et al. Laminectomy plus fusion versus laminectomy alone for lumbar spondylolisthesis. N Engl J Med 2016;374:1424–34. https://doi.org/10.1056/NEJMoa1508788.

[18] Försth P, Ólafsson G, Carlsson T, Frost A, Borgström F, Fritzell P, et al. A randomized, controlled trial of fusion surgery for lumbar spinal stenosis. N Engl J Med 2016;374:1413–23. https://doi.org/10.1056/NEJMoa1513721.

[19] 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. https://doi.org/10.3171/2014.4.SPINE14269.

[20] Groff MW, Dailey AT, Ghogawala Z, Resnick DK, Watters WC, Mummaneni PV, et al. Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 12: pedicle screw fixation as an adjunct to posterolateral fusion. J Neurosurg Spine 2014;21:75–8. https://doi.org/10.3171/2014.4.SPINE14277.

[21] Staartjes VE, Vergroesen P-PA, Zeilstra DJ, Schröder ML. Identifying subsets of patients with single-level degenerative disc disease for lumbar fusion: the value of prognostic tests in surgical decision making. Spine J 2018;18:558–66. https://doi.org/10.1016/j.spinee.2017.08.242.

[22] Schröder ML, Staartjes VE. Revisions for screw malposition and clinical outcomes after robot-guided lumbar fusion for spondylolisthesis. Neurosurg Focus 2017;42:E12. https://doi.org/10.3171/2017.3.FOCUS16534.

[23] Staartjes VE, Schröder ML. Effectiveness of a decision-making protocol for the surgical treatment of lumbar stenosis with grade 1 degenerative spondylolisthesis. World Neurosurg 2018;110:e355–61. https://doi.org/10.1016/j.wneu.2017.11.001.

[24] Schröder ML, de Wispelaere MP, Staartjes VE. Are patient-reported outcome measures biased by method of follow-up? Evaluating paper-based and digital follow-up after lumbar fusion surgery. Spine J 2018. https://doi.org/10.1016/j.spinee.2018.05.002.

[25] Lamers LM, Stalmeier PFM, McDonnell J, Krabbe PFM, van Busschbach JJ. Measuring the quality of life in economic evaluations: the Dutch EQ-5D tariff. Ned Tijdschr Geneeskd 2005;149:1574–8.

[26] Van Hooff ML, Spruit M, Fairbank JCT, van Limbeek J, Jacobs WCH. The Oswestry Disability Index (version 2.1a): validation of a Dutch language version. Spine 2015;40:E83–90. https://doi.org/10.1097/BRS.0000000000000683.

[27] Bursac Z, Gauss CH, Williams DK, Hosmer DW. Purposeful selection of variables in logistic regression. Source Code Biol Med 2008;3:17. https://doi.org/10.1186/1751-0473-3-17.

[28] Ondeck NT, Fu MC, Skrip LA, McLynn RP, Cui JJ, Basques BA, et al. Missing data treatments matter: an analysis of multiple imputation for anterior cervical discectomy and fusion procedures. Spine J n.d. doi:10.1016/j.spinee.2018.04.001.

[29] van Buuren S, Groothuis-Oudshoorn CGM. Mice: Multivariate Imputation by Chained Equations in R. J Stat Softw 2011:45. https://doi.org/10.18637/jss.v045.i03.

[30] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc 1995;57:289–300. https://doi.org/10.2307/2346101.

[31] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2018.

[32] Staartjes VE, de Wispelaere MP, Schröder ML. Improving recovery after elective degenerative spine surgery: 5-year experience with an enhanced recovery after surgery (ERAS) protocol. Neurosurg Focus 2019;46:E7. https://doi.org/10.3171/2019.1.FOCUS18646.

[33] Fritzell P, Hägg O, Wessberg P, Nordwall A, Swedish Lumbar Spine Study Group. 2001 Volvo Award Winner in Clinical Studies: lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish

Lumbar Spine Study Group. Spine 2001;26:2521–32. discussion 2532-2534.

[34] Glassman SD, Carreon LY, Djurasovic M, Dimar JR, Johnson JR, Puno RM, et al. Lumbar fusion outcomes stratified by specific diagnostic indication. Spine J Off 2009;9:13–21. https://doi.org/10.1016/j.spinee.2008.08.011.

[35] Asher AL, Devin CJ, McCutcheon B, Chotai S, Archer KR, Nian H, et al. Patient characteristics of smokers undergoing lumbar spine surgery: an analysis from the Quality Outcomes Database. J Neurosurg Spine 2017;27:661–9. https://doi.org/10.3171/2017.4.SPINE16984.

[36] Solberg TK, Sørlie A, Sjaavik K, Nygaard ØP, Ingebrigtsen T. Would loss to follow-up bias the outcome evaluation of patients operated for degenerative disorders of the lumbar spine? Acta Orthop 2011;82:56–63. https://doi.org/10.3109/17453674.2010.548024.

[37] Kahlert J, Gribsholt SB, Gammelager H, Dekkers OM, Luta G. Control of confounding in the analysis phase − an overview for clinicians. Clin Epidemiol 2017;9:195–204. https://doi.org/10.2147/CLEP.S129886.

[38] Yeh RW, Valsdottir LR, Yeh MW, Shen C, Kramer DB, Strom JB, et al. Parachute use to prevent death and major trauma when jumping from aircraft: randomized controlled trial. BMJ 2018;363:k5094. https://doi.org/10.1136/bmj.k5094.

[39] Staartjes VE, de Wispelaere MP, Vandertop WP, Schröder ML. Deep learning-based preoperative predictive analytics for patient-reported outcomes following lumbar diskectomy: feasibility of center-specific modeling. Spine J 2018. https://doi.org/10.1016/j.spinee.2018.11.009.

[40] Etter JF, Perneger TV. Analysis of non-response bias in a mailed health survey. J Clin Epidemiol 1997;50:1123–8.