



Characteristics and predicted outcome of patients lost to follow-up after degenerative lumbar spine surgery

C. Parai^{1,2} · O. Hägg² · C. Willers³ · B. Lind^{1,2} · H. Brisby^{1,4}

Received: 15 January 2020 / Revised: 22 May 2020 / Accepted: 1 July 2020 / Published online: 14 July 2020
© The Author(s) 2020

Abstract

Purpose The relatively large number of participants lost to follow-up (attrition) in spinal registers calls for studies that investigate the features of these individuals and their possible outcome. The aim was to explore the effect of attrition on patient-reported outcome in patients undergoing degenerative lumbar spine surgery. Three groups were studied: spinal stenosis (LSS), disc herniation (LDH) and degenerative disc disorder (DDD).

Methods Patients who underwent surgery for degenerative lumbar spine conditions during 2008–2012 according to registration in the Swespine national register were eligible for the study. Non-respondents were registered in Swespine prior to surgery, but not at follow-up. Swespine data were merged with hospital data from seven Swedish regions (65% of the population), Statistics Sweden, the National Patient Register and the Social Insurance Agency.

Baseline characteristics of non-respondents were described and compared to those of the respondents. Coefficients from regression analyses on PROM values for respondents were used to estimate the levels of PROM values for non-respondents, assuming the same effects of baseline characteristics for the two subgroups. Regression analyses were then conducted to identify variables associated with non-response. The results from the regression analyses were used to predict outcomes for patients with the characteristics of a non-respondent. Primary outcome variable in LSS and LDH was Global Assessment for leg pain, and in DDD, Global Assessment for back pain.

Results Age, sex, educational level, smoking, living alone, being born outside the EU, previous spine surgery and unexpected events before follow-up were factors that were significantly associated with non-response. Being born inside, the EU was important in all of the studied groups (LSS: OR 0.61 $p < 0.000$; LDH: OR 0.68 $p = 0.001$; DDD: OR 0.58 $p = 0.04$). For spinal stenosis patients, an unexpected event appeared particularly important (OR 3.40, $p = 0.000$). The predicted outcome of non-respondents was significantly worse than for respondents (LSS: 75.4% successful outcome vs. 78.7%; LDH: 53.9% vs. 58.2%; DDD: 62.7% vs. 67.5%. P-value in all groups = < 0.000).

Conclusion Attrition in Swespine cannot be ignored, as non-respondents were predicted to have worse outcome. The effect of attrition bias should always be considered when contemplating outcome recorded in a quality register with patients lost to follow-up.

Keywords Spine surgery · Quality register · Non-response to follow-up · Attrition bias

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00586-020-06528-y>) contains supplementary material, which is available to authorized users.

✉ C. Parai
catharina.parai@spinecenter.se

¹ Department of Orthopaedics, Institute of Clinical Sciences, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

Introduction

In all types of clinical studies, the ultimate goal is to have complete follow-up data of all included patients. Attrition by non-response at follow-up may threaten the validity of

² Spine Center Göteborg, Gothenburg, Sweden

³ Department of Neurobiology, Care Sciences and Society, Karolinska Institute, Stockholm, Sweden

⁴ Department of Orthopaedics, Sahlgrenska University Hospital, Gothenburg, Sweden

registers. Registers tend to demonstrate rather large proportions of non-respondents [1]. The Swespine national register systematically collects information on patients undergoing spine surgery. It has a good coverage and completeness, but also approximately 25% non-respondents at follow-up at one year [2]. Do these non-respondents jeopardize the validity of Swespine data?

The validity of spine registers in the context of patient dropout has been tested before in Denmark, Norway and Sweden [3–6]. In these studies, the outcome for non-respondents was assessed in either phone interviews or by postal questionnaires. The sample size varied between 318 and 633. The Norwegian study concluded that non-respondents in statistical terms could be treated as missing at random [4]. In the Danish study, there was no statistical difference in satisfactory results—measured with transition questions—between respondents and non-respondents at the 1-year follow-up [3]. Elkan et al. compared disc hernia patients registered in Swespine to a smaller cohort studied separately. Despite a missing data rate on the outcome variable level of more than 40% the Swespine cohort showed similar results to the separate group, which had a follow-up rate of 98%. Thus, it was concluded that patient-reported outcome data in Swespine was representative for large cohorts [5]. Endler et al. studied 351 patients registered in Swespine as operated for degenerative conditions in the lumbar spine. A total of 148 did not return the Swespine 2-year, or 5-year follow-up questionnaire but responses could be obtained from 115 of these through repeated postal or phone attempts. This “delayed response group” showed a significantly lower perceived change in back pain compared to the first-time respondents, but since other patient-reported outcomes (ODI, VAS and EQ-5D) were similar, the authors concluded that loss to follow-up did not hamper the interpretation of Swespine data [6].

Pushing participants, reluctant to respond to the initial follow-up, to respond through repeated attempts by mail or telephone may induce bias. The present paper seeks to avoid this matter by using a predictive model. In contrast to the previous studies, the current paper examines a large nation wide sample and a larger set of factors, through the linkage of Swespine data to several other registers, and outcome estimation of non-respondents is performed in an algorithmic multiple regression model. An increasing unwillingness to respond to surveys and also a lack of an accepted minimal response rate call for additional studies on registers that investigate the possible outcome of participants lost to follow-up [7].

The aims of this study were to explore if, and how, the outcome data reported by patients in a national quality register was affected by loss to follow-up and to assess if the earlier results could be replicated.

Patients and methods

Study population and data sources

All individuals who underwent degenerative lumbar spine surgery 2008–2012 (according to Swespine and the regional patient administrative system) and who answered the baseline questionnaire from Swespine were eligible for inclusion. Data were anonymized by Statistics Sweden and each observation received a unique identity number used for linkage. Observations were linked on patient level between data from Swespine and patient administrative systems for seven Swedish regions (inhabiting 65% of the Swedish population), administrative data from the National Patient Register and Statistics Sweden (socioeconomics) as well as the Swedish Social Insurance Agency (sick leave and disability pension). Non-respondents were defined as not having returned the one-year follow-up questionnaire.

All analyses were performed on three subgroups defined by a combination of diagnosis (ICD 10) and procedure (Nomesco classification of surgical procedures, NCSP) codes as described in Table 1. The subgroups were: Lumbar Spinal Stenosis (LSS) operated with decompression or decompression + fusion; Lumbar Disc Herniation (LDH) operated with endoscopic, microscopic or open discectomy; Degenerative Disc Disorder (DDD) operated with fusion (all types) or disc arthroplasty. Cases with pseudarthrosis as main diagnosis were excluded.

Patient characteristics and other factors used in the prediction models

The explanatory variables used in the regression analyses are listed in Table 2. They were defined in accordance with Swespine or based on diagnosis or procedure codes from the administrative systems.

The explanatory variables were included based on clinical relevance and to assess potential impact of any sociodemographic variables. Previous sick leave or disability pension were included as explanatory factors for the disc herniation and the DDD subgroups. These explanatory factors were not applied in regression analyses for spinal stenosis as only 45% of the subgroup sample was below 65 years old at surgery and hence below the general retirement age in Sweden at the time.

The variable unexpected event included ICD-codes for recurrent disc herniation, bleeding, leakage of CSF, restenosis, unspecified adverse event during surgery, infection, unspecified surgery-related damage, pain, cauda equina syndrome, wound rupture, complications associated with

Table 1 Definition of the subgroups LSS (Lumbar Spinal Stenosis), LDH (Lumbar Disc Herniation) and DDD (Degenerative Disc Disorder) by a combination of main diagnosis code and any of the intervention codes

Subgroup	Diagnosis	Main Diagnosis code (ICD-10)	Intervention	Procedure code (KVÅ)
LSS	Central canal stenosis	M480/M480K/	Nerve root decompression	ABC36
	Recess stenosis/	M480J	Decompression of the spinal canal	ABC56/ABC66
	Foraminal Stenosis	M4888/M4888K	Removal of facet joint cyst	NAM39
	Disc herniation*	M511/M511K	Fusion	NAG
	Facet joint cyst	M672	Discectomy	ABC07/ABC16/ABC26
LDH	Disc herniation	M511/M511K	Discectomy	ABC07/ABC16/ABC26
DDD	Degenerative Disc Disorder	M990	Disc prosthesis Fusion	NAB94/NAB96/ NAG

*Observations with disc herniation as diagnosis but combined with the procedure code for decompression of the spinal canal (i.e., not discectomy) were put in the LSS group

Table 2 Independent variables included in the regression analyses

Male sex
Age (< 40, 50–59, 60–69, > 70 years)
Highest educational level (Elementary school; High school; College/University)
Disposable income (1st quintile–5th quintile)
Social welfare
Living alone
Unemployed
Born inside EU
Smoke
Duration of leg pain (no leg pain; less than 3 months; 3–12 months; 1–2 yrs; more than 2 yrs)
Duration of back pain (no leg pain; less than 3 months; 3–12 months; 1–2 yrs; more than 2 yrs)
EQ-5D baseline (< 0, 0–0.19, 0.20–0.59, 0.60–1.00)
ODI baseline (< 20, 20–39, 40–59, 60–79, 80–100)
VAS leg pain baseline (< 20, 20–39, 40–59, 60–79, 80–100)
VAS back pain baseline (< 20, 20–39, 40–59, 60–79, 80–100)
Previous back surgery
Elixhauser comorbidity index
Sick-leave days year before surgery
Inpatient care the year prior to surgery
Unexpected adverse event

Categories are shown in the parentheses

vascular injury, mechanical complications associated with disc prosthesis, decubitus, as well as reintervention codes registered during the first twelve months after surgery.

Statistical analysis

Descriptive data are presented for baseline characteristics within each subgroup. In addition to crude values, a formal statistical analysis was performed for each variable between respondents and non-respondents, to recognize any statistically significant systematic differences between the groups (two-sided tests on a 5% significance level). The Chi-squared test was applied for dichotomous and ordinal variables, the

Kruskal–Wallis test was used for count variables, and the t-test was used for continuous variables. Descriptive statistics of data in the one-year follow-up questionnaire are also presented.

Two sets of regression analyses were computed. First, logistic regression analyses were performed with non-response as outcome, showing the degree of association between baseline variables and non-response at follow-up. Second, regression analyses were calculated, with patient-reported outcome measures (PROMs) at one-year follow-up as dependent variables. To enable comparisons of predicted PROM values for respondents and non-respondents, respectively, all explanatory variables included were used to predict the outcome of surgery,

for the individuals who answered the follow-up questionnaire as well as those who did not.

The primary outcome measure was Global Assessment (GA). It is an ordinal parameter of perceived pain (in back or leg) after surgery compared to preoperative pain, reported retrospectively. Successful outcome of surgery was dichotomized with GA = 1 (pain free) or GA = 2 (much better) referred to as GA_{SUCCESS}; and GA = 3 (somewhat better), GA = 4 (unchanged) or GA = 5 (worse) as not successful. GA = 0 (no pain before surgery) was excluded from analysis. The observed and predicted proportion of GA_{SUCCESS} (logistic regression), as well as values of the secondary measures Oswestry Disability Index (ODI), Visual Analogue Scale for back and leg pain, respectively (VAS_{BACK/LEG}), and Eurqol-5-dimensions index (EQ-5D) at the one-year follow-up (ordinary least squares regression) were calculated in the response group. All outcome measures have previously been well described [8].

Since the regression model excluded individuals with any independent variable(s) missing, it would not be accurate to compare the predicted outcome of the non-respondents to the observed outcome of all respondents. Instead, the comparison is made with respondents displaying complete sets of explanatory variables.

The overall quality of the logistic regression models is reported as the area under the curve (AUC) which indicates the ability of the model to distinguish patients with a successful outcome (GA_{SUCCESS}) from patients that do not. The AUC ranges from 0.5 to 1, where 1 indicates that the model correctly predicts 100% of the observations, while 0.5 means that the model is no better than chance. An AUC > 0.7 was considered an acceptable precision.

Adjusted R² values were calculated for the least squares regression models, which indicate the proportion of the variation in the outcome variables (i.e., final scores of the ODI, VAS_{BACK/LEG}, and EQ-5D) that can be explained by the model.

To further report on the accuracy of the models, the observed outcome (i.e., the outcome for respondents) was also compared to the outcome for (the same) respondents predicted by the models.

Ethical considerations

The Regional Ethical Review Board in Stockholm approved the study protocol (reference number 2013/1542-31/5).

Results

Baseline and follow-up characteristics, and health care consumption

All analyses were performed on each one of the three diagnosis groups. To avoid too large a number of tables, only

the LSS cohort is presented in the main document and the results for the remaining groups can be found as an online resource (ESM_1). The number of individuals in each group is found in the flowchart in Fig. 1.

Baseline variables, expressed as percentages or means, are given in Table 3. In all three diagnosis groups, many variables significantly differed between respondents and non-respondents. In absolute numbers, the differences were small. Health care consumption is presented in Table 4, showing that non-respondents consume a slightly higher amount of health care. Significant differences between respondents and non-respondents at the 95% confidence level are shown in bold. Descriptive statistics of one-year follow-up data are found in Table 5.

Variables predicting non-response

A younger age predicted non-response in all three diagnosis groups (for LSS, see Table 6), as did also male sex, and being born outside of the EU. Low disposable income, living alone and smoking predicted non-response in the LSS and LDH groups. A low level of education predicted non-response in the LDH and the DDD groups.

There were no significant differences in baseline PROM values, apart from EQ-5D in the LSS cohort which indicated that a low preoperative quality of life score increased the probability of non-response.

Among the variables expressing health care consumption, a higher comorbidity, previous spine surgery and unexpected events increased the probability of non-response.

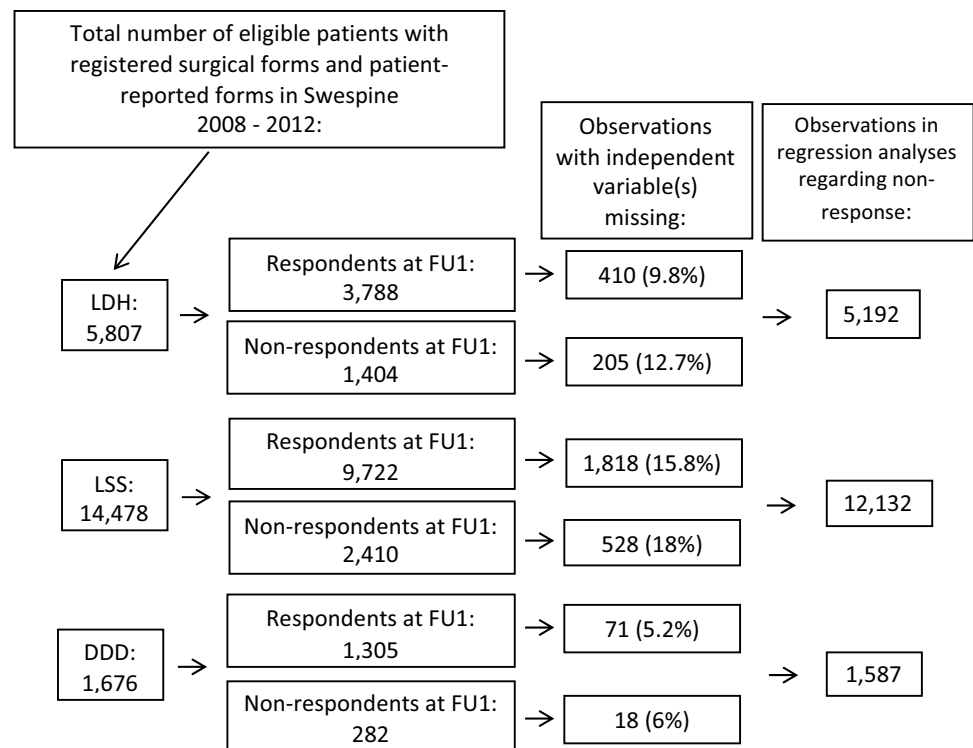
Variables predicting outcome

Variables predicting a successful outcome (according to the GA_{LEG}) for patients operated for spinal stenosis is presented in Table 7. The results for the LDH and DDD cohorts are shown in detail in the electronic supplementary material.

Male sex and living alone were factors associated with a decreased likelihood of successful outcome in the LSS group but not in the other groups. An age above 70 (reference age: < 40), smoking, a baseline of VAS_{BACK} > 40 (ref < 20) and a higher elixhauser comorbidity index decreased the probability in the LSS and the LDH groups. A higher level of education was associated with a higher probability of successful outcome in the LSS and LDH groups. Having the highest level of disposable income as well as being born inside the EU were variables associated with successful outcome in all three groups.

Having an unexpected event decreased the probability of successful outcome in the LSS and DDD cohorts. Previous back surgery decreased the likelihood in the DDD group.

Fig. 1 Flowchart showing the number of patients in each diagnosis group on the left, and on the right, the number of cases in the regression analyses, with non-respondent as dependent variable, after the exclusion of observations lacking one or several independent variables *LDH* Lumbar Disc Herniation cohort; *LSS* Lumbar Spinal Stenosis cohort; and *DDD* Degenerative Disc Disorder cohort



Predicted outcome

Predicted outcome for respondents and non-respondents measured with $GA_{\text{BACK/LEG}}$ is depicted in Fig. 2. The observed outcome for respondents 12 months after surgery and the predicted outcome for respondents and non-respondents, respectively, is presented in Table 8. The predicted outcome of non-respondents was significantly lower than both predicted and observed outcomes in all three diagnosis groups. The AUC, expressing the predictive ability of the model, was 0.73 for LDH; 0.69 for LSS; and 0.72 for DDD.

Significant regression equations were also found for the secondary outcome measures (ODI, $VAS_{\text{BACK/LEG}}$, and EQ-5D_{INDEX}). The quality of the models is summarized in Table 9. As was the case for the GA regression models, the observed outcome was highly consistent with the predicted outcome (in detail presented in Table 3 in Supplementary). The secondary measures displayed the same pattern as the GA, depicted in Fig. 3a-c, indicating that non-respondents are predicted to fare somewhat worse than respondents.

Discussion

The results in this study suggest that patients undergoing degenerative lumbar spine surgery and who are lost to follow-up in a national quality register demonstrate a somewhat worse outcome than follow-up respondents.

Similar findings have previously been described in a hip replacement cohort and in a longitudinal follow-up of rotator-cuff tears [9, 10]. The opposite conclusion that non-respondents do not influence the result and can be ignored in register studies has, however, been presented several times. In two such studies on spine register data, response was obtained by repeated attempts to reach dropouts, either by mail or by telephone interviews up to two years after the original follow-up [3, 4]. However, collecting data this way through interviews and telephone calls may induce bias [11]. Furthermore, all patients could not be reached in these studies. In the study by Hojmark et al. [3], the total number of non-respondents was only 48 and the non-response rate was 12%. Solberg et al. [4] studied 138 non-respondents, and the non-response rate was 22%. The current study analyzed a larger set of explanatory variables for a larger population, which might explain the inferential disparity. Elkan et al. [5] compared differences in outcome between the Swespine disc hernia population and a cohort of disc hernia patients from a single hospital. In the latter group, the response rates were considerably larger than in the Swespine register. After adjustments for covariates, the authors found no differences in outcome, between the two cohorts. However, individuals exhibiting characteristics, with a potential effect on non-response, such as previous spine surgery, linguistic difficulties and comorbidity were excluded. Further, only disc hernia patients were studied, while in the present study, larger differences were seen between non-respondents and respondents in the other diagnosis groups. The largest

Table 3 Baseline characteristics of patients operated for spinal stenosis

	Respondents			Non-respondents			Resp vs. non-respondents <i>p</i> value
	% or mean	SD	<i>n</i>	% or mean	SD	<i>n</i>	
Sex (males)	46		11,534	49		2933	0.010
Age	66	11	11,534	64	13	2933	0.000
Highest educational level			11,467			2900	0.004
Elementary school	30			31			
High school	44			46			
College/university	26			23			
Disposable income			11,528			2927	0.000
1st quintile	18			20			
2nd quintile	19			20			
3rd quintile	20			21			
4th quintile	21			20			
5th quintile	22			19			
Social welfare	2		11,530	4		2928	0.000
Living alone	39		11,530	43		2928	0.000
Unemployed	2		11,530	3		2928	0.007
Born inside EU	96		11,530	93		2928	0.000
Smoker	12		11,382	17		2893	0.000
Duration of leg pain			11,393			2893	0.839
No leg pain	3			3			
Less than 3 months	3			3			
3–12 months	26			26			
1–2 years	28			27			
More than 2 years	40			41			
Duration of back pain			11,360			2888	0.011
No back pain	6			4			
Less than 3 months	2			2			
3–12 months	19			19			
1–2 years	21			20			
More than 2 years	53			54			
EQ-5D baseline (0–1)	0.36	0.3	11,395	0.30	0.3	2871	0.000
ODI baseline (0–100)	43	16	11,075	46	16	2798	0.000
VAS _{LEG} baseline (0–100)	64	24	10,559	65	24	2653	0.002
VAS _{BACK} baseline (0–100)	56	26	10,514	59	26	2646	0.000
Previous back surgery	19		11,429	22		2891	0.000
Elixhauser comorbidity index (0–100)	51		11,534	62		2933	0.000
Sick-leave days year before surgery	55	99	5074	62	101	1479	0.017

Continuous variables are presented as means \pm SD and categorical data as percentage; EU the European Union

disparity in outcome was seen in the DDD group, suggesting that surgery for lumbar pain—as opposed to surgery for leg pain—is affected by losses to follow-up to a larger extent.

In a recent study [6], differences of similar size were found regarding Global Assessment for back pain, between respondents and non-respondents as reported in the present study. Still, the authors concluded that non-response could be ignored since both groups improved equally according to minimal important change (MIC) values in ODI,

NRS_{BACK/LEG} and EQ-5D. However, putting too much trust in score changes may, in the context of lost data, cause a relaxed attitude neglecting the importance of other outcome variables. Neither is reference to MIC theoretically obvious in this context of final outcome, since this value is based on calculation of the clinical importance of individual score changes.

A consequence of the differences between individuals lost to follow-up and respondents is that subgroups with certain

Table 4 Health care consumption in the spinal stenosis cohort

	Respondents			Non-respondents			Respondents vs. non-respondents <i>p</i> value
	% or mean	SD	n	% or mean	SD	n	
Inpatient care the year prior to surgery (number of days)	1	5	11,534	2	6	2933	0.000
Length of stay (number of perioperative days)	4	3	11,534	4	3	2933	0.984
Number of days of inpatient care first year related to adverse events (any complication)	7	11	583	7	9	351	0.475
Visits in specialized outpatient care first post-op year	3	4	11,534	4	5	2933	0.000
Physiotherapy (number of visits)	5	12	11,470	5	12	2913	0.099
Sickness benefit	29		11,534	32		2933	0.000
Unexpected event	3		11,534	8		2933	0.000

Continuous variables are presented as means ± SD and categorical data as percentage

Table 5 Descriptive statistics of observed 1-year follow-up data in the spinal stenosis cohort (respondents)

Outcome variable	% or mean	SD/CI	Count
GA _{LEG} (% self-assessing as pain-free or much better)	57.4	56.5–58.3	10,619
VAS _{LEG} (0–100)	33	31	10,598
ODI (0–100)	27	19	11,169
EQ-5D (0–1)	0.63	0.31	11,374
Satisfied with the results of surgery	65	63.8–65.5	11,205

Continuous variables are presented as means ± standard deviation (SD) and categorical data as percentage and confidence intervals (CI)

characteristics such as differences in socioeconomic status or adverse events are not fully recognized, leading to a decrease in external validity which threatens generalizability. This would mean that the register captures a selected subgroup that does not signify the whole target population. According to the current study, respondents are of older age, more often born inside of the EU, have a higher socioeconomic status, a better baseline health status and less adverse events, compared to non-respondents.

The differences between respondents and non-respondents in age, sex, smoking, level of education and income confirm the results of previous studies [3, 12–14]. Younger individuals were more likely to be non-respondents, which may reflect a lack of time because of work and family life.

Being born outside of the EU was significantly associated with non-response. This may reflect that cultural barriers negatively affect not just the outcome but may also be a factor of compliance with given information/instructions.

Non-respondents were more extensive consumers of health care. They made more visits to specialized outpatient care units during the first postoperative year, most likely because of remaining symptoms. As adverse events and previous back surgery were associated with non-response, it suggests that the register may fail to cover a subgroup that

Table 6 Regression analysis on the spinal stenosis cohort with non-response to the 1-year follow-up as dependent variable

Explanatory variable	OR	95% CI		<i>p</i> value
Male sex	1.15	1.04	1.26	0.004
Age				
< 40 yrs	[ref]			
40–49 yrs	0.70	0.52	0.88	0.005
50–59 yrs	0.54	0.42	0.69	0.000
60–69 yrs	0.42	0.33	0.53	0.000
> 70 yrs	0.41	0.32	0.52	0.000
Disposable income				
First quintile	[ref]			
Fifth quintile	0.82	0.70	0.95	0.011
Living alone	1.23	1.12	1.36	0.000
Born inside EU	0.61	0.50	0.75	0.000
Smoke	1.33	1.17	1.52	0.000
EQ-5D baseline				
< 0	[ref]			
0–0.19	0.79	0.67	0.94	0.009
Previous back surgery	1.14	1.02	1.28	0.021
Elixhauser comorbidity index	1.15	1.09	1.22	0.000
Unexpected event	3.40	2.81	4.11	0.000

An odds ratio (OR) >1 indicates a higher probability of being a non-respondent. An odds ratio <1 indicates a lower probability of being a non-respondent. Variables with non-significant results are left out in this table. The total list of variables is shown in table 2. EU the European Union

burdens the health care system and indirectly all taxpayers. It is important to capture this group of patients since; otherwise, an increase in adverse events over time or at specific centers/hospitals may not be recognized, but just reflected as a higher number of non-respondents. In addition, were these variables to have an influence on the outcome, it could be a threat to the internal validity of the register as it puts the inferences drawn from outcome data at risk. In other words,

Table 7 Regression analysis on the spinal stenosis cohort with the dichotomized patient self-assessment ($GA_{SUCCESS}$) as dependent variable

	OR	95% CI	<i>p</i> value
Sex (proportion men)	0.86	0.79 0.95	0.002
Age			
< 40 yrs	[ref]		
> 70 yrs	0.53	0.39 0.72	0.000
Highest educational level			
Elementary school	[ref]		
High school	1.16	1.05 1.29	0.005
College/university	1.59	1.39 1.79	0.000
Disposable income			
First quintile	[ref]		
Second quintile	1.19	1.02 1.38	0.021
Third quintile	1.22	1.05 1.41	0.008
Fifth quintile	1.31	1.13 1.52	0.001
Living alone	0.84	0.77 0.93	0.000
Born inside EU	1.96	1.52 2.52	0.000
Smoker	0.69	0.61 0.80	0.000
Duration of leg pain			
No leg pain	[ref]		
3–12 months	1.81	1.24 2.28	0.002
1–2 years	1.62	1.11 2.36	0.012
More than 2 years	1.28	1.11 2.36	0.012
Duration of back pain			
No back pain	[ref]		
Less than 3 months	1.83	1.11 3.00	0.017
EQ-5D baseline			
< 0	[ref]		
0–0.19	1.25	1.07 1.46	0.005
VAS _{BACK} baseline			
< 20	[ref]		
40–59	0.68	0.56 0.82	0.000
60–79	0.61	0.51 0.74	0.000
80–100	0.54	0.44 0.66	0.000
Previous back surgery	0.58	0.52 0.65	0.000
Elixhauser comorbidity index	0.87	0.82 0.92	0.000
Unexpected event	0.55	0.41 0.72	0.000

An odds ratio (OR) >1 indicates a higher probability of having a successful outcome (i.e., having responded as being “pain free” or “much better” on GA). An odds ratio <1 indicates a lower probability of being a non-respondent. Variables with non-significant results are left out in this table

the outcome of degenerative lumbar spine surgery may be misinterpreted and overrated.

Although not the main purpose of the current study, variables predicting outcome were presented in addition to variables predicting non-response (Table 7 for spinal stenosis cohort, and Tables 1e and 2e for disc hernia and degenerative disc disorder cohorts in the electronic supplementary

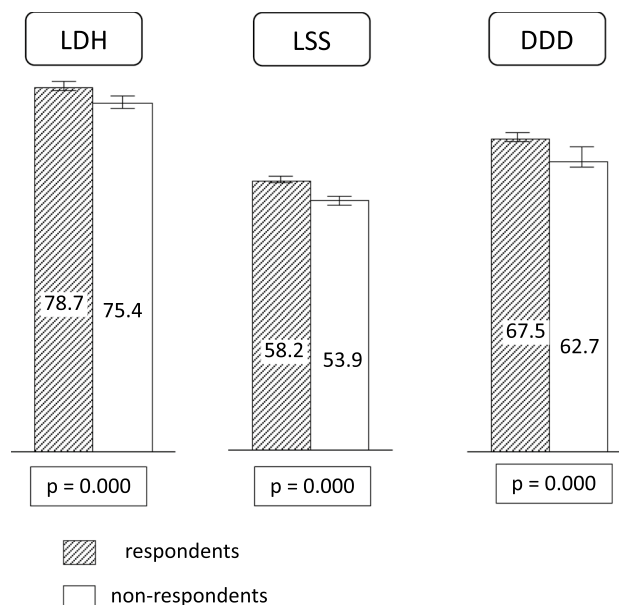


Fig. 2 Predicted proportions of a successful outcome according to Global Assessment (i.e., participants self-assessing as “pain free” or “much improved” at the one-year follow-up after the operation) for respondents and non-respondents, respectively. LDH Lumbar Disc Herniation (n=5,192); LSS Lumbar Spinal Stenosis (n=12,132); DDD Degenerative Disc Disorder (n=1,587); and Error bars indicate the 95% confidence intervals

material). Some risk factors predicted non-response *as well as* poor outcome (i.e., $GA_{SUCCESS} = no$). In the LSS cohort, these variables were male sex; living alone; born outside the EU; smoking; a higher Elixhauser comorbidity index; having had an adverse event during the first post-op year; and having had previous back surgery. In the LDH group, the corresponding variables were an educational level below university/college; born outside the EU; smoking; and comorbidity. In the DDD cohort, the single risk factor for both non-response and poor outcome was previous back surgery. These variables may be of particular interest in future case-mix adjusted outcome reports.

The current study had the advantage of access to several other registers and could thereby present a higher number of explanatory factors than previous studies. Although this adds some additional knowledge about non-response and its effects on the outcome, it is likely that there are still unknown factors. Furthermore, as presented in Fig. 1, there was a loss of observations also in the present study, due to missing values for the independent variables used for regression analyses.

Participants lost at the recruitment stage (i.e., patients not filling out the preoperative baseline questionnaire) were not included in this study, since they lack most baseline data necessary for analysis. Although we cannot exclude the possibility that they represent a subgroup with

Table 8 Percentage of patients self-assessing as having a successful outcome at the one-year follow up compared to the outcome for respondents and non-respondents, respectively, as predicted by the regression model

Diagn.group	GA _{SUCCESS}						
	Observed outcome (%)	95% Conf. Interv	Predicted outcome, respondents (%)	95% Conf. Interv	Predicted outcome non-respondents (%)	95% Conf. Interv	AUC
LDH	78.8	77.4–80.1	78.7	78.2–79.2	75.4	74.6–76.3	0.728
LSS	58.2	55.6–57.6	58.7	57.9–58.6	53.9	53.2–54.5	0.687
DDD	67.4	64.9–70.0	67.5	66.6–68.3	62.7	60.6–64.8	0.715

LDH Lumbar Disc Herniation cohort, *LSS* Lumbar Spinal Stenosis cohort, *DDD* Degenerative Disc Disorder cohort, *AUC* Area Under the Curve, which expresses the ability of the model to classify patients as either having a successful outcome or not having a successful outcome

Table 9 Predicted capability of least squares regression models with scores at the 1-year follow-up of four different PROMs as dependent variable

DV	LDH		LSS		DDD	
	<i>F</i> ratio*	Adjusted <i>R</i> ²	<i>F</i> ratio*	Adjusted <i>R</i> ²	<i>F</i> ratio*	Adjusted <i>R</i> ²
ODI	F(45, 3696)=34.44	0.29	F(45,9443)=101.97	0.32	F(45, 1249)=10.86	0.26
VAS _{LEG}	F(45,3687)=17,14	0.16	F(45,9326)=36.87	0.15	F(45, 1243)=9.16	0.22
VAS _{BACK}	F(45,3703)=23.20	0.21	F(45,9373)=55.20	0.21	F(45,1251)=5.88	0.14
EQ–5D	F(45,3716)=22.06	0.20	F(45,9563)=60.62	0.22	F(45,1252)=6.98	0.17

DV Dependent Variable, *LDH* Lumbar Disc Herniation cohort, *LSS* Lumbar Spinal Stenosis cohort, *DDD* Degenerative Disc Disorder cohort, *ODI* Oswestry Disability Index, *VAS* Visual Analogue Scale for back and leg pain respectively, *EQ–5D* EuroQol Five–dimensional quality of life questionnaire index, *F* is the ratio of the model mean square to the error mean square and is reported above in APA style

* indicates that all regression equations were significant at the $p < 0.000$ level

reluctance to fill in both baseline and follow-up questionnaires, it is our experience that the major cause of absent preoperative registration is logistic difficulties in certain hospitals.

Our data suggest that it is reasonable to assume that losses to follow-up in registers are missing not at random [16, 17]. What remains is to carefully reflect on the information that we do have to increase the awareness of what impact missing data might have and also suggest steps toward improvements that can increase the response to follow-up [11, 17].

The ability of the the logistic regression models to predict a successful outcome, expressed as AUCs, was 0.69–0.72. An acceptable level was considered 0.70 [15]. In a multiple imputation model, the uncertainty is incorporated in the model itself. The increased knowledge on the features of non-respondents presented in this study could be useful in future studies aiming to evaluate possible advantages of predicting outcome using multiple imputation.

An increased follow-up may be achieved by offering digital questionnaires in addition to the postal ones and to send

reminders via text messages and also to shorten the questionnaire. Information in other common languages about the purpose of data collection and also the preservation of anonymity provided in conjunction with the hospital stay and or on the register website may increase participation.

Conclusion

Outcome at the 1-year assessment of degenerative lumbar surgery is not expected to be the same for respondents and non-respondents. The pattern was seen for patients operated for disc herniation as well as for spinal stenosis and degenerative disc disorder. It is important to keep a closer eye on factors strongly predicting non-response and unsuccessful outcome such as unexpected adverse events—since the overall results otherwise may be overrated. The role of loss to follow-up in longitudinal register-based studies should not be underestimated.

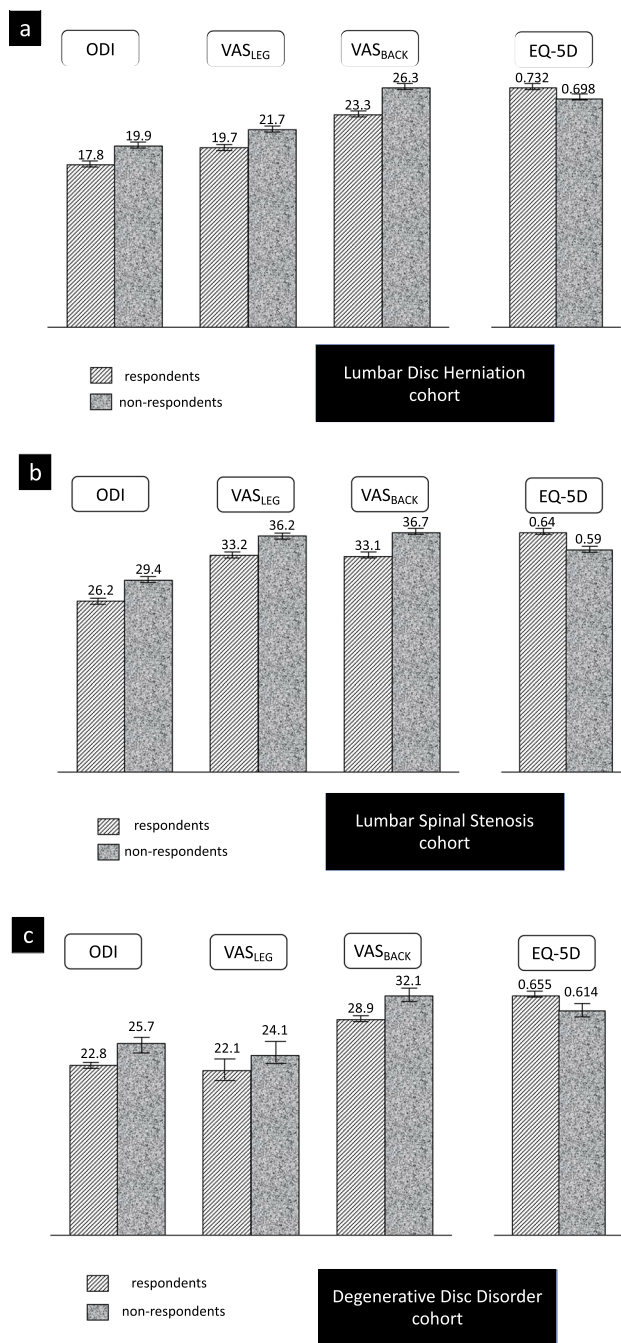


Fig. 3 a–c Predicted mean scores for respondents and non-respondents of the one-year follow-up questionnaires for three diagnosis groups. Error bars indicate 95% CI. LDH cohort $n = 5,192$; LSS cohort $n = 12,132$; DDD cohort $n = 1,587$; ODI Oswestry Disability Index; VAS Visual Analogue Scale for back and leg pain; EQ-5D Euroqol-five-dimensional index; Error bars indicate the 95% confidence intervals

Acknowledgements Open access funding provided by University of Gothenburg. The authors would like to thank the Neubergh foundation, the Region and the Regional Agreement on Medical Training and

Clinical Research (ALF-VG region 722931) and GHP Spine Center Göteborg for their financial support.

Compliance with ethical standards

Conflict of interest The 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. Lonne G, Fritzell P, Hagg O, Nordvall D, Gerdhem P, Lagerback T, Andersen M, Eiskjaer S, Gehrchen M, Jacobs W, van Hooff ML, Solberg TK (2019) Lumbar spinal stenosis: comparison of surgical practice variation and clinical outcome in three national spine registries. *Spine J Off J North Am Spine Soc.* <https://doi.org/10.1016/j.spinee.2018.05.028>
2. Stromqvist B, Fritzell P, Hagg O, Jonsson B, Sanden B (2013) Swespine: the Swedish spine register: the 2012 report. *Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc.* <https://doi.org/10.1007/s00586-013-2758-9>
3. Hojmark K, Stottrup C, Carreon L, Andersen MO (2016) Patient-reported outcome measures unbiased by loss of follow-up Single-center study based on Dane, Spine the Danish spine surgery registry. *Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc.* <https://doi.org/10.1007/s00586-015-4127-3>
4. Solberg TK, Sorlie A, Sjaavik K, Nygaard OP, Ingebrigtsen T (2011) Would loss to follow-up bias the outcome evaluation of patients operated for degenerative disorders of the lumbar spine? *Acta Orthop.* <https://doi.org/10.3109/17453674.2010.548024>
5. Elkan P, Lagerback T, Moller H, Gerdhem P (2018) Response rate does not affect patient-reported outcome after lumbar discectomy. *Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc.* <https://doi.org/10.1007/s00586-018-5541-0>
6. Endler P, Ekman P, Hellström F, Möller H, Gerdhem P (2019) Minor effect of loss to follow-up on outcome interpretation in the Swedish spine register. *Eur Spine J.* <https://doi.org/10.1007/s00586-019-06181-0>
7. Johnson TP, Wislar JS (2012) Response rates and nonresponse errors in surveys. *JAMA.* <https://doi.org/10.1001/jama.2012.3532>
8. Parai C, Hagg O, Lind B, Brisby H (2018) The value of patient global assessment in lumbar spine surgery: an evaluation based on more than 90,000 patients. *Eur Spine J Off Publ Eur Spine Soc Eur Spinal Deform Soc Eur Sect Cerv Spine Res Soc.* <https://doi.org/10.1007/s00586-017-5331-0>
9. Murray DW, Britton AR, Bulstrode CJ (1997) Loss to follow-up matters. *J bone joint surg Br.* <https://doi.org/10.1302/0301-620x.79b2.6975>
10. Norquist BM, Goldberg BA, Matsen FA 3rd (2000) Challenges in evaluating patients lost to follow-up in clinical studies of rotator cuff tears. *J bone joint surg Am.* <https://doi.org/10.2106/00004623-20006000-00010>

11. McKnight PE, McKnight KM, Sidani S, Figueredo AJ (2007) *Missing Data: A Gentle Introduction*. Guilford Press, New York
12. Juto H, Gartner Nilsson M, Moller M, Wennergren D, Morberg P (2017) Evaluating non-responders of a survey in the Swedish fracture register: no indication of different functional result. *BMC Musculoskelet Disord*. <https://doi.org/10.1186/s12891-017-1634-x>
13. Powers J, Tavener M, Graves A, Loxton D (2015) Loss to follow-up was used to estimate bias in a longitudinal study: a new approach. *J Clin Epidemiol*. <https://doi.org/10.1016/j.jclinepi.2015.01.010>
14. Jooste PL, Yach D, Steenkamp HJ, Botha JL, Rossouw JE (1990) Drop-out and newcomer bias in a community cardiovascular follow-up study. *Int J Epidemiol*. <https://doi.org/10.1093/ije/19.2.284>
15. DW Hosmer Jr S Lemeshow RX Sturdivant 2013 *Applied Logistic Regression* 398 John Wiley & Sons Hoboken
16. Kristman V, Manno M, Cote P (2004) Loss to follow-up in cohort studies: how much is too much? *Eur J Epidemiol*. <https://doi.org/10.1023/b:ejep.0000036568.02655.f8>
17. Graham JW (2009) Missing data analysis: making it work in the real world. *Annu Rev Psychol*. <https://doi.org/10.1146/annurev.psych.58.110405.085530>

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