USASP
US ASSOCIATION FOR THE STUDY OF PAIN

PUBLISHED BY
ELSEVIER

The Journal of Pain, Vol 23, No 8 (August), 2022: pp 1318–1342
Available online at www.jpain.org and www.sciencedirect.com

Systematic Review and Meta-Analysis of Predictors of Return to Work After Spinal Surgery for Chronic Low Back and Leg Pain

Monika Halicka,* Rui Duarte,† Sharon Catherall,‡ Michelle Maden,† Michaela Coetsee,* Martin Wilby,§ and Christopher Brown*

*Department of Psychological Sciences, University of Liverpool, Liverpool, UK
†Liverpool Reviews & Implementation Group (LRiG), University of Liverpool, Liverpool, UK
‡Public Health Policy and Systems / LRiG, University of Liverpool, Liverpool, UK
§Department of Neurosurgery, The Walton Centre NHS Foundation Trust, Liverpool, UK

Abstract: **Spinal surgeries to treat chronic low back pain (CLBP) have variable success rates, and despite the significant personal and socioeconomic implications, we lack consensus for prognostic factors. This systematic review and meta-analysis evaluated the evidence for preoperative predictors of return to work (RTW) after spinal surgery for CLBP. We searched electronic databases and references (January 1984 to March 2021), screened 2,622 unique citations, and included 8 reports (5 low and 3 high risk-of-bias) which involved adults with ≥3 months duration of CLBP with/without leg pain undergoing first elective lumbar surgery with RTW assessed ≥3 months later. Narrative synthesis and meta-analysis where possible found that individuals less likely to RTW were older (odds ratio [OR] = .58; 95% confidence interval [CI]: 0.46–0.72), not working before surgery, had longer sick leave (OR = .95; 95% CI: 0.93–0.97), higher physical workload, legal representation (OR = .61; 95% CI: 0.53–0.71), psychiatric comorbidities and depression (moderate quality-of-evidence, QoE), and longer CLBP duration and opioid use (low QoE), independent of potential confounders. Low quality and small number of studies limit our confidence in other associations. In conclusion, RTW after spinal surgery for CLBP likely depends on sociodemographic and affective psychological factors, and potentially also on symptom duration and opioid use.**

Perspective: ***This systematic review and meta-analysis synthesizes and evaluates existing evidence for preoperative predictors of return to work after spinal surgery for chronic low back pain. Demonstrated associations between return to work and sociodemographic, health-related, and psychological factors can inform clinical decision-making and guide further research.***

Key words: *Chronic low back pain, spinal surgery, lumbar spine, predictors, return to work.*

Introduction

Low back pain affects 40% of people at some point in their lives[33] and is the leading cause of years lived with disability in the world.[46,55] The personal and socioeconomic burden is particularly high for persistent or recurring low back pain, estimated to affect 60% of people a year after an initial acute episode.[27] When conservative treatment fails to reduce painful symptoms

Supported by Translational Research Access Programme (TRAP), Faculty of Health & Life Sciences, University of Liverpool, UK. The funder approved the objective of this review but had no role in data collection and synthesis, decision to publish, or preparation of the manuscript.
The authors have no conflicts of interest to declare.
Address reprint requests to Monika Halicka, PhD, Department of Psychological Sciences, University of Liverpool, Eleanor Rathbone Building, Bedford Street South, Liverpool L69 7ZA, UK. E-mail: mon.halicka@gmail.com
1526-5900/$36.00

https://doi.org/10.1016/j.jpain.2022.02.003

and improve function, surgical interventions can be considered to address the underlying spinal pathology.[15,37] The previous decade has seen more than a twofold increase in the number of lumbar spine surgeries performed in the UK and US.[45,54] While spinal surgery is one of the most invasive and expensive approaches to manage chronic low back pain (CLBP), its benefit is often suboptimal. A minimal clinically important reduction in pain intensity is reported only by 60% of patients undergoing first lumbar spine surgery[30,54,57,58] and long-term healthcare costs of CLBP that persists or recurs after surgery are 50% higher than those for patients without ongoing pain. Nearly 80% of the total cost of low back pain can be attributed to indirect costs associated with work absence and productivity loss.[14]

Indeed, return to work (RTW) rates after lumbar spine surgery are highly variable (3−100%),[35] partly due to patient heterogeneity. Identifying which factors can reliably predict the likelihood of RTW could inform development of clinical prediction models, facilitate managing patients' expectations, which are strongly associated with work participation outcomes in CLBP more generally,[29] and help to determine the most beneficial and cost-effective course of treatment based on individual socioeconomic, health-related, and psychological characteristics. For example, preoperative cognitive-behavioral therapy has been shown to improve outcomes of spinal fusion surgery.[48]

Two previous systematic reviews suggested that patients who are older, female, have comorbidities and longer symptom duration, who are not working before surgery, have higher physical workload, occupational mental stress, passive pain coping, or depression are less likely to RTW after surgery for radiculopathy due to lumbar disc herniation.[35,43] These results were almost all based on single studies and could not be pooled for meta-analysis. Insufficient or conflicting evidence regarding prognosis of spinal surgery outcomes precludes development of clinical guidelines on relevant predictors.[15,37] Furthermore, the existing reviews were restricted to a specific population, while prognostic factors for RTW may be common across different spinal pathologies and surgical measures and thus provide stronger evidence for the examined associations. Importantly, chances of recovery and response to treatments decrease with longer duration of low back pain.[13,18,27,42,52] While this may affect the prognosis of RTW,[50] so far there has been no evidence synthesis focusing on surgical candidates with CLBP. Therefore, we aimed to identify and evaluate preoperative predictors of RTW after spinal surgery for CLBP with or without radicular pain.

Methods

This review was prospectively registered on PROSPERO (CRD42020180845) before commencing the screening stage. Our methodology and reporting followed the general principles for conducting reviews in health care outlined in the Centre for Reviews and Dissemination (CRD) guidance[10] and the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA).[44] The current article is 1 of the 2 planned review papers on the predictors of spinal surgery outcomes. The current paper focuses on predictors of return to work, and the second paper concerns predictors of patient-reported pain and disability outcomes.[59]

Search Strategy

Electronic databases MEDLINE, EMBASE, PsycINFO, CINAHL, and Cochrane Central Register of Controlled trials (CENTRAL) were searched on April 08, 2020 and updated March 29, 2021, following a search strategy developed in collaboration with an information specialist (M.M.; for full search strategy, see Supplementary Text S1). Search results were exported to EndNote Library and deduplicated. A manual search for additional primary studies was also conducted through the reference lists of relevant systematic reviews and included studies.[8]

Eligibility Criteria

Eligibility criteria are outlined according to a modified PICOTS format for reviews of prognostic studies (Population, Index and Comparator prognostic factors, Outcomes, Timing, Setting).[47]

Population consisted of adults with CLBP, defined as pain lasting or recurring for ≥3 months,[36] with or without lumbar radicular pain, defined as pain radiating to the leg due to nerve root compression.[51] Only patients undergoing primary lumbar or lumbosacral spine surgery, without history of previous spinal surgery, were eligible. We excluded spinal pathologies such as cancer / tumor, inflammatory disease, infection, or trauma, as well as spinal cord stimulator implantation, injections, radiofrequency, chemical interventions, and studies which investigated the impact of pre- or postoperative interventions. There was no requirement for all patients to be working before surgery.

Index prognostic factors included variables assessed prior to surgery and investigated for their potential ability to predict RTW after surgery. Radiographic, genetic, and any postoperative or intraoperative predictors were beyond the scope of the current review and were excluded, unless used as potential confounders. We examined both adjusted prognostic effects from multivariate analyses, that is, independent effects of particular index prognostic factors over and above other (ie, comparator) prognostic factors, and unadjusted prognostic effects from univariate analyses, if reported.[47] There is no consensus on a minimum set of *comparator prognostic factors*, therefore, this criterion did not restrict study eligibility.

Eligible studies reported *outcomes* including RTW, as an objective measure of functional recovery. Any definition of RTW was acceptable. The *timing* of outcome assessment was ≥3 months after surgery, with no upper limit. Study *settings* such as spinal surgery sites or registries of operated patients were eligible for inclusion.

Study *designs* had to allow investigation of associations between preoperative prognostic factors and postoperative RTW outcomes, and could include

randomized or nonrandomized controlled studies, cohort, case-control, or registry-based studies, with prospective or retrospective designs. Case reports and case series were excluded as providing only low level of evidence in prognosis research.[2,56]

Publication formats included original peer-reviewed studies published between January 1984 (when diagnosis of spinal pathologies and surgical treatments could be informed by magnetic resonance imaging)[40] and March 2021 in English language. Conference abstracts or unpublished research were not included.

Study Selection

Titles and abstracts of deduplicated records, and then full texts, were screened against the eligibility criteria by 2 independent reviewers (M.H. and R.D.) using a piloted form. Disagreements were resolved by consensus or seeking opinion from a third reviewer (M.W.) where necessary. If eligibility was unclear based on the abstract, the report was moved to full text screening. In case of uncertainty about eligibility based on the information available in the full text, supplement, and any related publications, study authors were contacted to request additional details. The selection process is illustrated in a PRISMA diagram[44] (Fig 1) and presents the primary reasons for exclusion recorded as the first category for which eligibility criteria were not met with certainty.

Data Extraction

A data extraction form was developed based on the Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prognostic Factor Studies (CHARMS-PF)[47] and piloted by 2 independent reviewers (M.H. and S.C.; for the final template, see Supplementary Table S1). Each reviewer then extracted the data from half of the included reports and verified the accuracy of the data extracted by the other reviewer from the remaining reports. Any inconsistencies were resolved by discussion and consensus.

The following information was extracted from the included reports: study design, setting, method and time of participant recruitment and eligibility criteria, baseline participant characteristics, type of surgery, sample size estimation, number of included participants and participation rate, number, definition, method and time of measurement of index and comparator prognostic factors and outcomes, method of handling continuous factors and missing data, response rate, reasons for loss to follow-up and characteristics of lost participants, analysis methods, assumptions, methods and criteria for selecting predictors for and during multivariable modelling, adjusted and unadjusted effect estimates where available, signs of selective reporting, appropriateness of interpretation and discussion.[47]

The desired common effect estimate for binary RTW outcomes was odds ratio (OR) with confidence interval (CI) or standard error (SE), however, depending on the analysis type, regression coefficients or mean differences were also extracted. Where the desired effect estimates were not provided, we calculated these from available data (eg, 2 × 2 tables or means and standard deviations) or transformed the reported estimates using effect size calculators.[1,38,39]

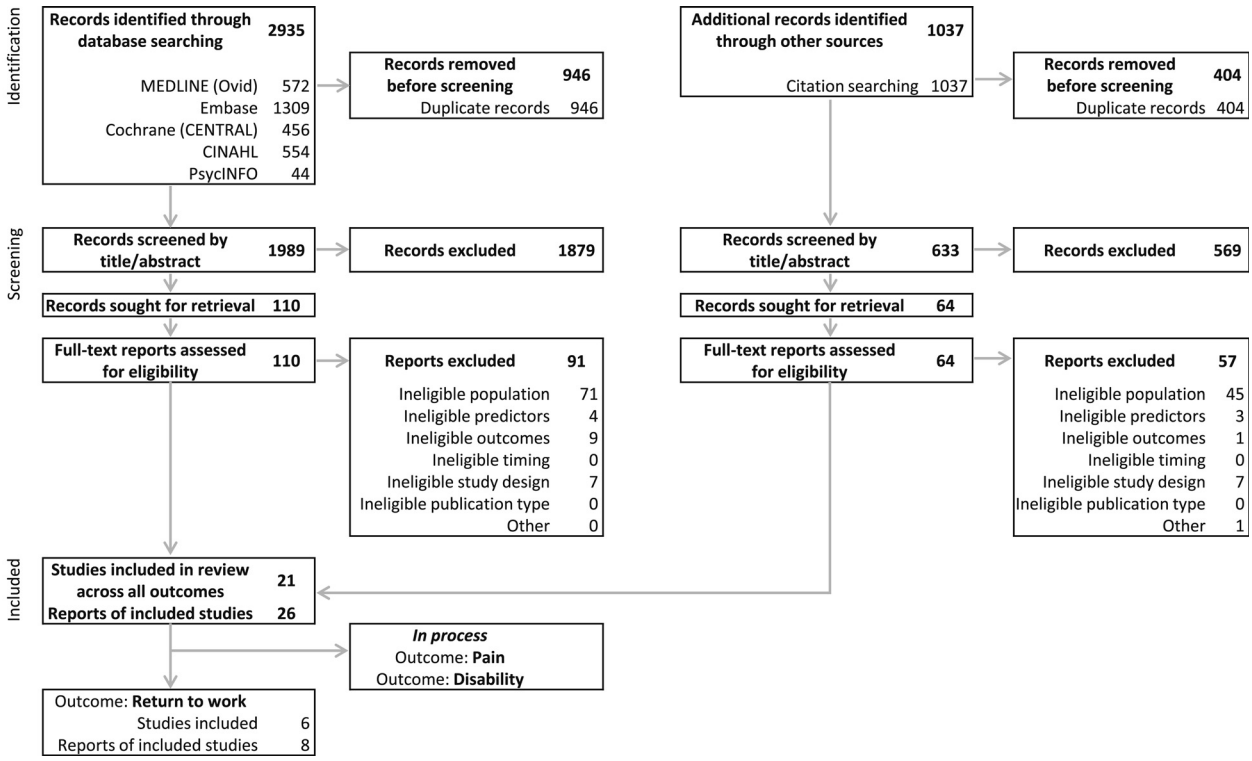

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram.[44]

Risk of Bias Assessment

Risk of bias (RoB) of included reports was assessed using the Quality in Prognosis Studies (QUIPS) tool[21,26,28] in 6 domains, each rated as being at high, moderate, or low RoB: study participation, study attrition, prognostic factor measurement, outcome measurement, study confounding, and statistical analysis and reporting. RoB ratings were guided by prompting items suggested by Grooten et al[21] and adapted to the current review question (for the final QUIPS template, see Supplementary Table S2). Two reviewers (M.H. and S.C.) independently assessed half of the included reports, and each checked the judgements of the other reviewer for agreement or domains with unclear RoB. Any inconsistencies were resolved by discussion and consensus. Finally, the overall 'low' or 'high' RoB ratings were assigned to each report, reflecting whether all QUIPS domains had low-moderate RoB ('low'), or if ≥1 domain had high RoB ('high').[9] Results of QUIPS assessment were considered during synthesis of the results and grading the quality of evidence.

Data Synthesis

Where there was sufficient and appropriate data for quantitative synthesis, we performed meta-analysis of the effects of index prognostic factors on RTW. Analyses were conducted in R software using *meta* package.[7,25] A minimum of 2 studies reporting data on the same predictor were required to pool the results. Precalculated prognostic effect estimates (ORs with SE or 95% CI) of included factors were pooled using random effects generic inverse variance. Random effects models were preferred as a degree of clinical and methodological heterogeneity was present among the included studies. The results of each meta-analysis were presented in a forest plot as pooled estimate of the average effect of the predictor with 95% CI and estimates of statistical heterogeneity (I^2 and tau^2). Tau^2 was estimated using Paule-Mandel procedure as a recommended method when outcomes are binary and the number of pooled studies is small.[6,53] Substantial between-study heterogeneity was indicated by I^2 >75% or 95% CI of tau^2 not including 0 and $P < .05$.[31] Meta-analyses were performed for the following candidate predictors deemed suitable for quantitative synthesis: age, gender, marital status, work status, duration of sick leave, worker's compensation, legal representation, income, disability, and symptom duration. For a large proportion of these associations, only unadjusted effect estimates could be pooled. Due to small number of studies contributing to each meta-analysis, we did not conduct sensitivity analyses that would include only low RoB studies.

Meta-analyses were not conducted for other associations due to methodological heterogeneity or insufficient data. Specifically, it was not possible to quantitatively combine evidence from adjusted and unadjusted analyses, and with different definitions of predictors (eg, timescales of opioid use and other analgesics). Insufficient results reporting (missing effect estimates and / or their precision), and only single studies contributing evidence for certain associations (eg, personality traits / disorders, pain-related psychological factors), further prevented meta-analyses.

Since quantitative synthesis was only possible for some predictors, we included a tabular summary of all extracted adjusted and unadjusted associations with RTW, with a narrative synthesis of the findings. This synthesis presented the number of studies investigating each association of interest, the magnitude and direction of any reported effects, and assessed the consistency of available evidence between contributing studies. The evaluation of the findings also accounted for the RoB judgements at the study and outcome level.

Grading of Evidence

Grading of Recommendations, Assessment, Development and Evaluations (GRADE) framework[22] adapted for reviews of prognostic studies[34] guided the evaluation of the quality of evidence for associations between predictor categories and RTW. Two reviewers (M.H. and M.C.) collaboratively determined the overall quality of evidence as high, moderate, low, or very low, respectively reflecting high, moderate, limited, and very limited confidence that the true effect lies close to the observed estimate of that effect.[22]

The starting quality of evidence depended on the phase of investigation, specifically, it was high for phase-3 and phase-2 studies (assessing prognostic pathways / mechanisms, and independent associations between hypothesized predictor and outcome, respectively), and moderate for phase-1 studies (exploring potential associations between prognostic factors and outcome).[34] If any hypotheses-driven studies presented only unadjusted analyses or investigated broad classes of multiple predictors, they were classified as phase-1. Following the recommendations by Huguet et al,[34] the starting quality rating could be downgraded for severe study limitations (due to dominant evidence from high RoB studies or unadjusted analyses), clinically meaningful inconsistency in the effect estimates, indirectness (eg, if it was not possible to verify chronic pain status of all included participants), imprecision (due to inadequate sample size or insufficient results reporting), and publication bias (if certain relationship was investigated in <4 studies); and upgraded if moderate (OR ≥2.5 or ≤0.4) or large (OR ≥4.25 or ≤0.24) effect size, or 'dose' effect (where higher levels of the predictor would lead to greater effect sizes) were present, resulting in an overall quality rating.

Results

Study Selection

We screened 2,622 unique records and their flow through the selection process is illustrated in Figure 1. A list of excluded articles with reasons is provided in Supplementary Text S2. The most common reason for exclusion was ineligible population where patients without chronic pain or those who underwent previous spinal

surgery were included. We were not able to confirm the chronic pain status of all participants in 3 studies, however, available average symptom duration data suggested that they were likely eligible.[4,16,42] This uncertainty was taken into account during the evaluation of the quality of evidence. Studies without any information about symptom duration, or surrogate indicators such as duration of conservative treatment, were excluded. Reports that assessed surgical and nonsurgical cohorts were included if data specific to eligible surgical group could be extracted.[24,41] Overall, 8 eligible reports of 6 studies assessing predictors of RTW after surgery were included in the current review.

Study Characteristics

Table 1 provides an overview of study characteristics. Five reports were based on prospective studies (2 single-[5,49] and 3 multicenter[16,23,24]). Since 2 of them analyzed the same participants from the Swedish Lumbar Spine Study, we considered them to reflect a single cohort.[23,24] Three reports included retrospective cohorts from the Ohio Bureau of Workers' Compensation.[4,41,42] They used the data from overlapping time periods, therefore, 2 reports that focused on fusion surgery were considered to correspond to a single cohort,[4,41] while another report on single-level discectomy was considered to reflect a separate cohort.[42] Three of the included studies were based in the United States and the remaining ones in the Netherlands, Sweden, and Switzerland. The most common pathologies were disc herniation, degenerative disc disease and spondylosis, and surgical measures included fusion and discectomy. All participants were employed prior to surgery, except for Haag's study (86% employed),[23,24] and unclear employment status in Anderson's 2006 study as *not working* could also refer to being on sick leave.[5] Most studies defined RTW outcome as return to any (part- / full-time) work, 2 required RTW to be sustained for ≥6 months, and 1 considered only RTW in full capacity relative to before symptom onset. Follow-up duration ranged from 6 to 36 months, and RTW rates were between 26% and 90% across included studies.

Risk of Bias in Included Studies

The reviewers made consistent RoB judgements on 83% of QUIPS domain ratings, with Cohen's kappa = .70, 95% CI: 0.52 to 0.88, suggesting substantial agreement,[3,12] before they reached final consensus. RoB ratings in each domain across all reports are illustrated in Figure 2. Overall, 5 reports were judged to have low, and 3 high RoB (where at least 1 QUIPS domain was rated as high).[9]

Most reports had low RoB in study participation, prognostic factor and outcome measurement, study confounding, and statistical analysis and reporting domains. Moderate RoB in study attrition domain in almost all included reports resulted from missing information about reasons for and / or characteristics of participants lost to follow-up, despite overall high attrition rates (min. 81% and >90% in most studies). RoB in 2 studies that retrospectively recruited only participants with complete follow-up[4,42] was also rated as moderate. Serious study limitations were identified in study participation, confounding, and statistical analysis and reporting domains in single studies. Specifically, high RoB in these domains was due to insufficient information regarding baseline characteristics, recruitment setting, and number of screened participants;[5] not adjusting for any confounders;[24] and missing information about statistical assumptions and partial results reporting.[49]

Results of Syntheses

Predictors of RTW were examined in 6 unique patient cohorts (4616 participants in total). Table 2 details the results for the associations analyzed in each included study. Where both adjusted and unadjusted effects were available, we only present the synthesis of independent associations below. Synthesis of unadjusted relationships can be found in Supplementary Text S3.

Sociodemographic Predictors of RTW

Five studies (6 reports) investigated the associations between demographic and socioeconomic factors and RTW in 5 cohorts including a total of 4,574 unique patients.

Older *age* was an independent negative predictor of RTW in 2 low RoB studies[4,23] but not in another high RoB study reporting adjusted analysis.[5] This observed inconsistency does not appear to be explained by sample characteristics, specific RTW criteria, duration of follow-up, or whether age was analyzed as a continuous or categorical factor; however, this discrepancy is unlikely to be clinically meaningful, and low RoB studies that did find a significant effect of age consistently reported its negative direction (including in unadjusted analyses). The pooled adjusted effect of age (defined as a categorical predictor, >48 or >50 years) supported a significant negative association (Fig 3A).

The independent effect of preoperative *work status* on RTW was examined in 1 phase-2 high RoB study, which found a large positive effect of working before surgery.[5]

Three low RoB studies assessed the prognostic value of being on *sick leave* and / or its duration for RTW outcomes in adjusted analyses. Inability to continue working during the week before surgery negatively predicted RTW (small effect).[4] Longer duration of sick leave was also a negative predictor of RTW in 3 studies (small effects),[16,23,41] however, its independent effect in 1 study did not reach statistical significance.[16] Pooled adjusted OR of sick leave duration indicated a significant small negative effect on RTW (Fig 3B).

A relationship between having *legal representation* and RTW was assessed in 2 low RoB studies, both reporting small-moderate negative effects in adjusted analyses.[4,41,42] Pooled adjusted OR was consistent with a significant small negative effect of having legal representation on RTW (Fig 3C).

Table 1. Study Characteristics

Study ID, setting	Study type	Population				Index & comparator prognostic factors*	Outcomes*	Timing
		Inclusion / exclusion criteria†	Sample characteristics	Diagnosis (pathology)	Surgery (type, levels)	Measures	Measure (RTW rate)	Follow-up (resp. rate)
Anderson 2006;[5] Orthopedic Surgery and Rehabilitation Department, University Hospital (US)	Prospective cohort	Exclusion: significant psychosocial abnormalities upon psychological assessment	N = 106; Age <48 and >48; Duration ≥6 mo; 47% working	Discogenic LBP	Anterior lumbar interbody fusion	Work status at time of surgery; Smoking; Gender; Worker's compensation; Age; Baseline pain; Baseline disability; Levels fused; Cage type	Working, including homemaker / students resuming duties and return to different work / with restrictions (67%)	24 mo (81%)
Anderson 2015‡;[4] Ohio Bureau of Workers' Compensation (US)	Retrospective cohort (registry-based)	Inclusion: fusion between 1993 and 2013, ≥3 y follow-up available; Exclusion: smoking, permanent disability	N = 2,799; Age M = 43, SD = 10; 72% male; Duration M = 25.5, SD = 20.64 mo; 100% employed, 12% working week before surgery	57% disc herniation, 51% lumbar sprain, 28% degenerative disc disease, 17% spondylolisthesis, 13% radiculopathy, 10% spinal stenosis, 7% spondylosis, 2% spondylolysis, <1% scoliosis	Fusion (68% 1-level, 32% multilevel; 34% instrumented, 16% device, 41% instrumented and device, 9% noninstrumented)	Depression; Working before surgery; Age; Opioid use before surgery; Legal representation; Spinal pathology; Lumbar surgery after index fusion; Type of fusion; Graft type	Return to work within 2 y sustained for ≥6 mo (32%)	36 mo (100%)
Den Boer 2006;[16] 4 Hospitals (the Netherlands)	Prospective cohort, post hoc analysis	Inclusion: failure of conservative treatment, employed before the symptom onset; Exclusion: physical comorbidity that could interfere with postoperative rehabilitation	N = 182; Age 19 to 61, M = 41; 59% male; Duration M = 9.25, SD = 7 mo; 100% employed, on sick leave M = 14, SD = 20 wk	Disc herniation with radiculopathy	Disc surgery	Education level; Baseline disability; Neurological deficits; Postoperative pain intensity; Fear of movement; Pain coping; Physical workload; Job satisfaction; Duration of sick leave; Age; Gender; Baseline pain; Symptom onset; Medication; Duration of current pain episode; Outcome expectations	Return to work in full capacity, relative to before symptom onset (78%)	6 mo (91%)
Hagg 2003‡;[24] 19 Orthopedic Departments (Sweden)	RCT, post hoc analysis	Inclusion: severe chronic LBP, ≥2 y duration, back pain more severe than leg pain, no radiculopathy, ≥1 y of sick leave / equivalent disability / failed conservative treatment, ≥7/10 Function and Working Disability Score; Exclusion: psychiatric illness, spondylolisthesis, spinal stenosis, painful and disabling arthritic hip joints	N = 201; Age 25 to 64, M = 43, SD = 8; 49% male; Duration 24 −408, M = 94.08, SD = 81.84 mo; 86% employed, 91% on sick leave, M = 25, SD = 30 mo	Degenerative spondylosis	Posterolateral fusion (67% instrumented, 33% noninstrumented)	Pain drawing	Working part- / full-time (38%)	24 mo (91%)
Hagg 2003‖;[23] 19 Orthopedic Departments (Sweden)						Age; Gender; Occupation; Work status; Marital status; Comorbidity; Worker's compensation; Duration of back pain; Duration of sick leave; Smoking; Personality traits (neuroticism, aggressiveness, social introversion, impulsiveness); Personality disorders; Depressive symptoms; Pain		

(continued on next page)

Table 1. Continued

Study ID, setting	Study type	Population				Index & comparator prognostic factors*	Outcomes*	Timing
		Inclusion / exclusion criteria†	Sample characteristics	Diagnosis (pathology)	Surgery (type, levels)	Measures	Measure (RTW rate)	Follow-up (resp. rate)
						behavior; Back pain; Leg pain; Disability; General Function Score; Finger to floor distance; Pain on flexion / extension; Painful segment; Motor function; Tendon reflexes; Sensation		
Nguyen 2011[c];[41] Ohio Bureau of Workers' Compensation (US)	Retrospective case-control (registry-based)	Inclusion: injury between 1999 and 2001; Exclusion: surgery after 2003, injuries to body parts other than lumbar spine, head trauma, pregnancy	N = 725; Age 18 to 70, M = 40, SD = 9; 72% male; Duration M = 21, SD = 12 mo; 100% employed, on sick leave M = 337, SD = 277 d	48% disc herniation, 24% disc degeneration, 12% radiculopathy, 9% spondylolisthesis, 6% spinal stenosis	Fusion (68% 1-level, 32% multilevel)	Age; BMI; Time off presurgery; Time from injury to surgery; Diagnosis; Education; Gender; Legal representation; Marital status; Smoking; Wages; Complications; Daily morphine; Reoperation	Return to part- / full-time work 2 y after surgery (26%)	24, M = 57 mo (99%)
O'Donnell 2018;[42] Ohio Bureau of Workers' Compensation (US)	Retrospective cohort (registry-based)	Inclusion: ≥3 y follow-up; Exclusion: spondylolisthesis, spinal deformity, epidural hematomas / abscess, smoking, multilevel surgery	N = 1,286; Age 18 to 80, M = 40; 76% male; Duration M = 13, SD = 13 mo; 100% employed	Disc herniation	1-level discectomy	Opioid use before surgery; Time from injury to surgery; Legal representation; Psychiatric comorbidity; Household income	Return to same / different work within 2 y sustained for ≥6 mo (55%)	36 mo (89%)
Schade 1999;[49] Orthopedic and Neurosurgery Departments (Switzerland)	Prospective case-control, post hoc analysis	Inclusion: 6 to 8 wk of failed conservative treatment, availability for clinical and MRI examination before surgery; Exclusion: rapid progressive motor deficit, cauda equina syndrome, no Swiss residentship	N = 46; Age 20 to 50, M = 35; 74% male; Duration 46% 3 to 6, 26% 6 to 12, 28% >12 mo; 100% employed	Disc herniation with radiculopathy	Discectomy	Job-related resignation; Baseline pain/disability; Depression; Occupational mental stress; Vitality; Anxiety; General job satisfaction	Return to any work (90%)	24, range 23–30 mo (91%)

Abbreviations: BMI, Body Mass Index; LBP, Low Back Pain; M, mean; MRI, Magnetic Resonance Imaging; RCT, Randomized Controlled Trial; RTW, return to work; SD, standard deviation.
*Only eligible index prognostic factors or those included in eligible analyses as comparator prognostic factors are listed; only eligible outcomes are listed.
†Listed only inclusion/exclusion criteria that were not covered by the eligibility criteria for the systematic review.
‡Studies based on the same or overlapping populations: Anderson 2015 and Nguyen 2011 (the same database, partially overlapping dates); Hagg 2003 and Hagg 2003 (the same cohorts from the Swedish Lumbar Spine Study).

Study ID	Bias Domains						
	1. Study Participation	2. Study Attrition	3. Prognostic Factor Measurement	4. Outcome Measurement	5. Study Confounding	6. Statistical Analysis and Reporting	Overall Assessment of RoB
Anderson 2006 [3]	H	M	L	M	L	M	High
Anderson 2015 [2]	M	M	M	L	M	L	Low
Den Boer 2006 [6]	L	L	L	L	L	L	Low
Hagg 2003 [22]	L	M	L	L	H	M	High
Hagg 2003 [21]	L	M	M	L	M	M	Low
Nguyen 2011 [39]	M	M	M	L	L	L	Low
O'Donnell 2018 [40]	L	M	L	L	M	L	Low
Schade 1999 [47]	L	M	M	M	L	H	High

Figure 2. Risk of bias judgements based on Quality in Prognosis Studies (QUIPS).[28] Overall assessment of RoB: Low = all domains low or moderate; High = one or more domains high.[9] Abbreviations: L, low; M, moderate; H, high risk of bias.

One low RoB study investigated the effect of *physical workload*, showing a small independent negative association between higher physical workload score and RTW.[16]

An independent association between *income* and RTW was found in 2 low RoB studies, where participants with higher weekly wages and household income had higher odds of RTW after surgery (small effects).[41,42] However, the pooled adjusted OR was not statistically significant and there was some indication of heterogeneity (Fig 3D).

We found no significant independent associations with RTW for other sociodemographic factors, that is, *gender* (1 high RoB study),[5] general *education* level (1 low RoB study),[16] or *workers' compensation* status (1 high RoB study).[5] *Marital status* was only investigated in unadjusted analyses in 2 low RoB studies,[23,41] with pooled OR indicating no significant effect (Fig 3E).

Health-Related Predictors of RTW

Six studies (7 reports) examined the associations between health- and symptom-related factors and RTW outcomes in 6 cohorts including a total of 4617 unique patients.

One low RoB study assessed whether type of *spinal pathology* affected RTW, showing a small independent effect of having spondylosis on lower odds of RTW.[4] Notably, several pathologies were considered in the adjusted analysis, yet only spondylosis was significant and reported.

Independent prognostic value of *symptom duration* for RTW was examined in 1 low RoB study. Longer time from injury to surgery was a significant negative predictor of RTW after surgery,[42] and this association was supported by a pooled unadjusted OR from 3 other low RoB studies[16,23,41] (see Supplementary Text S3).

Two low RoB studies, including a phase-2 study, investigated the independent effect of *analgesics use* on RTW. Both found significant associations indicating that using (compared to not using) opioids before surgery,[42] and using them for longer than (compared to less than) a year,[4] was associated with lower odds of RTW (small effects).

No significant independent associations with RTW were found for other health-related factors, including preoperative *pain* (1 high RoB study),[5] *disability* (1 low and 2 high RoB studies),[5,16,49] sensory and motor *neurological signs* (1 low RoB study),[16] and *smoking* status (1 high RoB study).[5] Effects of acute *symptom onset* (1 low RoB study),[16] *comorbidities* (1 low RoB study),[23] and *BMI* (1 low RoB study)[41] were only investigated in unadjusted analyses, showing no significant associations with RTW after surgery.

Psychological Predictors of RTW

Five studies (6 reports) assessed the relationships between psychological factors and RTW in 5 cohorts including 4510 unique patients in total.

One low RoB study assessed the association between *psychiatric comorbidity* and RTW. Participants with any psychiatric comorbidity (including affective disorders and schizophrenia) had lower odds of RTW after surgery in adjusted analysis (moderate effect size).[42]

Two studies (1 low, 1 high RoB)[4,49] examined independent effects of *depression* on RTW. Participants with a clinical diagnosis of depression in 1 phase-2 study,[4] and those with higher depression scores in another study,[49] had lower odds of RTW relative to those without depression diagnosis or lower scores in adjusted analyses (moderate and small effect sizes, respectively).

The prognostic value of *pain coping* was assessed in 1 low RoB study, suggesting that passive pain coping was associated with reduced work capacity in adjusted analysis (small effect).[16]

The same low RoB study assessed the effect of *kinesiophobia* (fear of movement-related pain) on RTW,

Table 2. Effects of Prognostic Factors on Return to Work

Study ID	Prognostic factor	Outcome		Sample size	Analysis	Effect estimates
	Measure, definition	Measure, definition	Follow-up (mo)	N analyzed	Method, adjusted for factors* / unadjusted	Estimate (95% CI)†
	Demographic					
	Age					
Anderson 2006[5]	Age <48 y versus >48 y	Work status at follow-up, working versus not working	24, M = 30	106	I. Logistic regression adjusted for: baseline work status (working), smoking (yes), gender (NR), worker's compensation (yes), age >48, baseline pain (0–10 VAS), baseline disability (RMDQ), levels fused (single), cage type (BAK)	I. Age <48 y: OR = 1.41 (0.37–5.33), $P = .61$; Age >48 y: OR = .39 (0.07–2.22), $P = .29$
Anderson 2015[4]	Age >50 versus ≤50 y	Return to work within 2 y after surgery sustained for ≥6 mo versus failure to return to work	36	2799 (612 >50, 2187 ≤50)	I. Logistic regression adjusted for: depression (yes), working within a week before surgery (yes), opioids use for >1 y before surgery (yes), legal representation (yes), spinal pathology (spondylosis), another lumbar surgery after index fusion (yes), type of fusion (ALIF, PLF + PLIF), graft type (allograft)	I. OR = .58 (0.46–0.73), $P < .001$
Hagg 2003[23]	Age (y)	Work status at follow-up, working (part- or fulltime) versus not working	24	169	I. Stepwise forward regression analysis adjusted for: current sick leave (mo)	I. Beta = -.05; OR = .95 (0.91–0.99)
				201 (74 working, 127 not working)	II. Independent t-test (unadjusted)	II. Working M = 40, SD = 7.3, not working M = 44, SD = 8.5; $P = .001$; §d = -.50 (-0.79 to 0.20)
Nguyen 2011[41]	Age (y)	Return to work (part- or full-time) 2 y after surgery versus failure to return to work in any capacity	24	725	I. Logistic regression (unadjusted)	I. OR = 1.00 (0.98–1.02), $P = .81$
Den Boer 2006[16]	Age (y)	Full return to work (100%) versus reduced work capacity (<100%; self-reported % of work capacity compared to before symptom onset)	6	182 (141 full, 41 reduced capacity)	I. Logistic regression (unadjusted)	I. b = -.03, $P > .05$; full M = 40.4, SD = 9.1, reduced M = 43, SD = 9.7; §OR = .6 (0.32–1.13)
	Gender					
Anderson 2006[5]	NR	Work status at follow-up, working versus not working	24, M = 30	106	I. Logistic regression adjusted for: baseline work status (working), smoking (yes), worker's compensation (yes), age <48, age >48, baseline pain (0–10 VAS), baseline disability (RMDQ), levels fused (single), cage type (BAK)	I. OR = .56 (0.17–1.82), $P = .34$
Nguyen 2011[41]	Male versus female	Return to work (part- or full-time) 2 y after surgery versus failure to return to work in any capacity	24	725	I. Logistic regression (unadjusted)	I. OR = .90 (0.60–1.32), $P = .57$
Hagg 2003[23]	Male versus Female	Work status at follow-up, working (part- or fulltime) versus not working	24	201 (99 male, 102 female)	I. 2 × 2 frequency table (Chi-squared test) (unadjusted)	I. 44 female, 30 male were working at follow-up; $P > .05$; §OR = .57 (0.32–1.02)
Den Boer 2006[16]	Male versus female	Full return to work (100%) versus reduced work capacity (<100%; self-reported % of work capacity compared to before symptom onset)	6	182 (107 male, 75 female)	I. Logistic regression (unadjusted)	I. b = -.04, $P > .05$; §OR = .96

(continued on next page)

Table 2. Continued

Study ID	*Prognostic factor*	*Outcome*		*Sample size*	*Analysis*	*Effect estimates*
	Measure, definition	Measure, definition	Follow-up (*mo*)	N analyzed	Method, adjusted for factors* / unadjusted	Estimate (95% CI)†
	Marital status					
Hagg 2003[23]	Married or cohabitating versus not married	Work status at follow-up, working (part- or fulltime) versus not working	24	200 (159 married, 41 not married)	I. 2 × 2 frequency table (Chi-squared test) (unadjusted)	I. 60 married, 14 not married were working at follow-up; $P > .05$; §OR = 1.19 (0.58–2.45)
Nguyen 2011[41]	Divorced / married / widowed versus single	Return to work (part- or full-time) 2 y after surgery versus failure to return to work in any capacity	24	725	I. Logistic regression (unadjusted)	I. $P = .49$; divorced OR = 1.26 (0.70–2.26), $P = .44$; married OR = 1.37 (0.89–2.11), $P = .16$; widowed OR = .61 (0.07–5.62), $P = .66$
	Education					
Nguyen 2011[41]	Education level: high-school / no college / college or higher versus no high-school	Return to work (part- or full-time) 2 y after surgery versus failure to return to work in any capacity	24	725	I. Logistic regression (unadjusted)	I. $P = .06$; high-school OR = 1.99 (1.12–3.51), $P = .02$; no college OR = 1.44 (0.69–3.00), $P = .34$; college/higher OR = 3.13 (1.06–9.28), $P = .04$
Den Boer 2006[16]	Educational level: primary (<8 y) versus secondary (8–14 y) versus tertiary (>14 y)	Full return to work (100%) versus reduced work capacity (<100%; self-reported % of work capacity compared to before symptom onset)	6	182 (51 primary, 86 secondary, 45 tertiary)	I. Logistic regression adjusted for: fear of movement (TSK), passive pain coping (Pain Coping Inventory), negative outcome expectancies, physical workload, job satisfaction, duration of sick leave (wk), gender (male), age, baseline disability (RMDQ) and pain (0–100 VAS), 3 d postoperative pain (0–100 VAS), analgesics week before surgery (types), neurologic deficits (none / sensory or motor, both), duration (wk) and onset (acute) of symptoms; II. Logistic regression (unadjusted)	I. b = -.02, SE = .14, $P = .91$, OR = 1.02 II. b = -.18, $P < .05$; §OR = .84
	Work status					
Anderson 2006[5]	Working (including home working and studies) versus not working	Work status at follow-up, working versus not working	24, M = 30	105 (49 working, 56 not working)	I. Logistic regression adjusted for: smoking (yes), gender (NR), worker's compensation (yes), age <48, age >48, baseline pain (0–10 VAS), baseline disability (RMDQ), levels fused (single), cage type (BAK); II. 2 × 2 frequency table (unadjusted)	I. OR = 10.5 (2.64–41.4), $P = .001$; II. 45 working, 24 not working were working at follow-up; §OR = 15.00 (4.74–47.44)
Hagg 2003[23]	Employed versus unemployed	Work status at follow-up, working (part- or fulltime) versus not working	24	201 (173 employed, 28 unemployed)	I. 2 × 2 frequency table (Chi-squared test) (unadjusted)	I. 64 employed, 10 unemployed were working at follow-up; $P > .05$; §OR = 1.06 (0.46–2.43)
	Sick leave					
Nguyen 2011[41]	Number of days off prior to surgery	Return to work (part- or full-time) 2 y after surgery versus failure to return to work in any capacity	24	725	I. Stepwise logistic regression adjusted for: weekly wages, complications (early major systemic / implant /late spinal / neurologic / wound / none), legal representation (no), daily morphine, reoperation (yes); II. Logistic regression (unadjusted)	I. OR = .94 (0.92–0.97), $P < .001$; II. OR = .93 (0.91–0.95), $P < .001$
Den Boer 2006[16]	Duration of sick leave for the current pain episode (wk)	Full return to work (100%) versus reduced work capacity (<100%; self-reported % of work capacity compared to before symptom onset)	6	182 (141 full, 41 reduced capacity)	I. Logistic regression adjusted for: fear of movement (TSK), passive pain coping (Pain Coping Inventory), negative outcome expectancies, physical workload, job satisfaction, gender (male), age, education level (primary / secondary / tertiary), baseline disability (RMDQ) and pain (0–100 VAS), 3 d postoperative pain (0–100	I. b = -.23, SE = .19, $P = .24$, OR = .79; II. b = -.46, $P < .01$; full M = 12.5, SD = 22.1, reduced M = 17.5, SD = 12.5; §OR = .64 (0.34–1.21)

(*continued on next page*)

Table 2. Continued

Study ID	Prognostic factor	Outcome		Sample size	Analysis	Effect estimates
	Measure, definition	Measure, definition	Follow-up (mo)	N analyzed	Method, adjusted for factors* / unadjusted	Estimate (95% CI)†
					VAS), analgesics week before surgery (types), neurologic deficits (none / sensory or motor, both), duration (wk) and onset (acute) of symptoms; II. Logistic regression (unadjusted)	
Hagg 2003[23]	Duration of current sick leave (mo)	Work status at follow-up, working (part- or fulltime) versus not working	24	169	I. Stepwise forward regression analysis adjusted for: age (y)	I. Beta = -.036; OR = .96 (0.94–0.99)
				179 (59 working, 120 not working)	II. Independent t-test (unadjusted)	II. Working M = 14, SD = 14, not working M = 31, SD = 35; $P < .001$; §d = .57 (0.25–0.89)
	On sick leave versus not on sick leave			201 (183 on sick leave, 18 not on sick leave)	III. 2 × 2 frequency table (Chi-squared test) (unadjusted)	III. 64 on sick leave, 12 not on sick leave were working at follow-up; $P = .01$; §OR = .27 (0.09–0.74)
Anderson 2015[4]	Inability versus ability to continue working within the same week as surgery	Return to work within 2 y after surgery sustained for ≥6 mo versus failure to return to work	36	2799 (349 working, 2450 not working)	I. Logistic regression adjusted for: depression (yes), age (>50 y), opioids use for >1 y before surgery (yes), legal representation (yes), spinal pathology (spondylosis), another lumbar surgery after index fusion (yes), type of fusion (ALIF, PLF + PLIF), graft type (allograft)	I. OR = .47 (0.36–0.60), $P < .001$
	Worker's compensation					
Anderson 2006[5]	Compensation claim versus no compensation	Work status at follow-up, working versus not working	24 (mean 30)	106 (50 compensation, 56 no compensation)	I. Logistic regression adjusted for: baseline work status (working), smoking (yes), gender (NR), age <48, age >48, baseline pain (0–10 VAS), baseline disability (RMDQ), levels fused (single), cage type (BAK); II. 2 × 2 frequency table (unadjusted)	I. OR = .69 (0.20–2.32), $P = .54$; II. 28 compensation, 42 no compensation were working at follow-up; §OR = .42 (0.19–0.97)
Hagg 2003[23]	Compensation claim versus no compensation	Work status at follow-up, working (part- or fulltime) versus not working	24	192 (105 compensation, 87 no compensation)	I. 2 × 2 frequency table (Chi-squared test) (unadjusted)	I. 30 compensation, 39 no compensation were working at follow-up; $P = .035$; §OR = .49 (0.27–0.90)
	Legal representation					
Nguyen 2011[41]	Having versus not having legal representation	Return to work (part- or full-time) 2 y after surgery versus failure to return to work in any capacity	24	725	I. Stepwise logistic regression adjusted for: weekly wages, complications (early major systemic / implant / late spinal / neurologic / wound / none), days off pre-surgery, daily morphine, reoperation (yes); II. Logistic regression (unadjusted)	I. OR = .29 (0.13–0.63), $P = .002$; II. OR = .25 (0.14–0.46), $P < .001$
Anderson 2015[4]	Having versus not having legal representation	Return to work within 2 y after surgery sustained for ≥6 mo versus failure to return to work	36	2799 (1821 with, 978 without legal representation)	I. Logistic regression adjusted for: depression (yes), working within a week before surgery (yes), age (>50), opioid use for >1 y before surgery (yes), spinal pathology (spondylosis), another lumbar surgery after index fusion (yes), type of fusion (ALIF, PLF + PLIF), graft type (allograft)	I. OR = .64 (0.53–0.77), $P < .001$
O'Donnell 2018[42]	Having versus not having legal representation	Return to (same/different) work within 2 y sustained for ≥6 mo versus no sustained return to work	36	1286 (700 with, 586 without legal representation)	I. Stepwise logistic regression adjusted for: opioid use before surgery (yes), time from injury to surgery (mo), psychiatric comorbidity (yes), mean household income ($)	I. OR = .57 (0.44–0.73), $P < .01$
	Physical workload					

(continued on next page)

Table 2. Continued

Study ID	Prognostic factor	Outcome		Sample size	Analysis	Effect estimates
	Measure, definition	Measure, definition	Follow-up (*mo*)	N analyzed	Method, adjusted for factors* / unadjusted	Estimate (95% CI)†
Hagg 2003[23]	Having a heavy job versus not heavy job	Work status at follow-up, working (part- or fulltime) versus not working	24	181 (55 heavy, 126 not heavy job)	I. 2 × 2 frequency table (Chi-squared test) (unadjusted)	I. 26 heavy job, 47 not heavy job were working at follow-up; $P > .05$; §OR = 1.51 (0.80–2.87)
Den Boer 2006[16]	Physical workload score (5–20; Questionnaires on Musculoskeletal Load and Health Complaints)	Full return to work (100%) versus reduced work capacity (<100%; self-reported % of work capacity compared to before symptom onset)	6	182 (141 full, 41 reduced capacity)	I. Logistic regression adjusted for: fear of movement (TSK), passive pain coping (Pain Coping Inventory), negative outcome expectancies, job satisfaction, duration of sick leave (wk), gender (male), age, education level (primary / secondary / tertiary), baseline disability (RMDQ) and pain (0–100 VAS), 3 d postoperative pain (0–100 VAS), analgesics week before surgery (types), neurologic deficits (none / sensory or motor, both), duration (wk) and onset (acute) of symptoms; II. Logistic regression (unadjusted)	I. b = -.17, SE = .06, $P < .001$, OR = .84; II. b = -.12, $P < .01$; full M = 9, SD = 3.8, reduced M = 10.9, SD = 3.7; §OR = .40 (0.21–0.76)
	Income					
Nguyen 2011[41]	Weekly wages ($)	Return to work (part- or full-time) 2 y after surgery versus failure to return to work in any capacity	24	725	I. Stepwise logistic regression adjusted for: complications (early major systemic / implant / late spinal / neurologic / wound / none), days off presurgery, legal representation (no), daily morphine, reoperation (yes); II. Logistic regression (unadjusted)	I. OR = 1.12 (1.03–1.21), $P = .008$; II. OR = 1.13 (1.05–1.20), $P < 0.001$
O'Donnell 2018[42]	Mean household income ($, estimated based on zip codes and US census data)	Return to (same/different) work within 2 y sustained for ≥6 mo versus no sustained return to work	36	1286	I. Stepwise logistic regression adjusted for: opioid use before surgery (yes), time from injury to surgery (mo), legal representation (yes), psychiatric comorbidity (yes)	I. OR = 1.01 (1.00–1.02), $P < .01$
	Health-related					
	Pain					
Anderson 2006[5]	VAS (0–10) pain intensity	Work status at follow-up, working versus not working	24 (mean 30)	106	I. Logistic regression adjusted for: baseline work status (working), smoking (yes), gender (NR), worker's compensation (yes), age <48, age >48, baseline pain (0–10 VAS), levels fused (single), cage type (BAK)	I. OR = .82 (0.60–1.12), $P = .22$
Den Boer 2006[16]	VAS (1–100) average intensity of back and leg pain intensity over the past week	Full return to work (100%) versus reduced work capacity (<100%; self-reported % of work capacity compared to before symptom onset)	6	182 (141 full, 41 reduced capacity)	I. Logistic regression (unadjusted)	I. b = -.01, $P > .05$; full M = 45, SD = 20.9, reduced M = 52.5, SD = 21.7; §OR = .52 (0.28–0.99)
Hagg 2003[23]	VAS (0–100) average of worst, least, and current back pain intensity	Work status at follow-up, working (part- or fulltime) versus not working	24	199 (73 working, 126 not working)	I. Independent t-test (unadjusted)	I. Working M = 64, SD = 13.2, not working M = 64, SD = 15; $P > .05$; §d = .00 (-0.29 to 0.29)
	VAS (0–100) average of worst, least, and current leg pain intensity			199 (73 working, 126 not working)		I. Working M = 35, SD = 23.6, not working M = 35, SD = 26.5; $P > .05$; §d = .00 (-0.29 to 0.29)
	Pain on flexion versus no pain on flexion			197 (126 pain, 71 no pain)	II. 2 × 2 frequency table (Chi-squared test) (unadjusted)	II. 50 pain, 21 no pain on flexion were working at follow-up, $P > .05$; §OR = 1.56 (0.84–2.89)

(*continued on next page*)

Table 2. Continued

Study ID	Prognostic factor	Outcome		Sample size	Analysis	Effect estimates
	Measure, definition	Measure, definition	Follow-up (mo)	N analyzed	Method, adjusted for factors* / unadjusted	Estimate (95% CI)†
	Pain on extension versus no pain on extension			196 (141 pain, 55 no pain)		II. 52 pain, 20 no pain on extension were working at follow-up, $P > .05$; §OR = 1.00 (0.52–1.91)
	Positive versus negative springing test (localization of painful segment by palpation)			198 (182 positive, 16 negative)		II. 46 positive, 6 negative were working at follow-up, $P > .05$; §OR = .54 (0.19–1.58)
	Disability					
Anderson 2006[5]	RMDQ score (0–24)	Work status at follow-up, working versus not working	24 (mean 30)	106	I. Logistic regression adjusted for: baseline work status (working), gender (NR), worker's compensation (yes), age <48, age >48, baseline pain (0–10 VAS), baseline disability (RMDQ), levels fused (single), cage type (BAK)	I. OR = .93 (0.81–1.07), $P = .30$
Den Boer 2006[16]	RMDQ score (0–24)	Full return to work (100%) versus reduced work capacity (<100%; self-reported % of work capacity compared to before symptom onset)	6	182 (141 full, 41 reduced capacity)	I. Logistic regression adjusted for: fear of movement (TSK), passive pain coping (Pain Coping Inventory), negative outcome expectancies, physical workload, job satisfaction, duration of sick leave (wk), gender (male), age, education level (primary / secondary / tertiary), baseline pain (0–100 VAS), 3 d postoperative pain (0–100 VAS), analgesics week before surgery (types), neurologic deficits (none / sensory or motor, both), duration (wk) and onset (acute) of symptoms; II. Logistic regression (unadjusted)	I. b = -.01, SE = .07, $P = .83$, OR = .98; II. b = -.10, $P < .05$; full M = 14.8, SD = 3.2, reduced M = 16.3, SD = 4.2; §OR = .45 (0.24–0.86)
Hagg 2003[23]	General Function Score (0–100)	Work status at follow-up, working (part- or fulltime) versus not working	24	201 (74 working, 127 not working)	I. Independent t-test (unadjusted)	I. Working M = 46, SD = 14.3, not working M = 49, SD = 16.9; $P > .05$; §d = -.19 (-0.47 to 0.10)
	ODI score (0–100)					I. Working M = 46, SD = 9.7, not working M = 48, SD = 12.3; $P > .05$; §d = -.18 (-0.47 to 0.11)
	Finger-floor distance at maximum flexion (cm)			178 (66 working, 112 not working)		I. Working M = 29, SD = 17, not working M = 31, SD = 19; $P > .05$; §d = -.11 (-0.41–0.20)
Shade 1999[49]	Preoperative level of pain and RMDQ score (0–24)‡	Return to any work	24 (23–30)	42	I. Stepwise regression adjusted for: depression, occupational mental stress	I. Beta = -.26, T = 1.87, $P < .10$; §OR = .77
	Neurological signs					
Hagg 2003[23]	Normal versus abnormal reflexes	Work status at follow-up, working (part- or fulltime) versus not working	24	202 (171 normal, 31 abnormal)	I. 2 × 2 frequency table (Chi-squared test) (unadjusted)	I. 65 normal, 9 abnormal reflexes were working at follow-up, $P > .05$; §OR = 1.43 (0.62–3.27)
	Normal versus abnormal sensation			177 (132 normal, 45 abnormal)		I. 48 normal, 15 abnormal sensation were working at follow-up, $P > .05$; §OR = 1.14 (0.56–2.34)
	Normal versus abnormal motor function			201 (175 normal, 326 abnormal)		I. 67 normal, 8 abnormal motor function were working at follow-up, $P > .05$; OR = 1.36 (0.56–3.31)

(*continued on next page*)

Table 2. Continued

Study ID	Prognostic factor	Outcome		Sample size	Analysis	Effect estimates
	Measure, definition	Measure, definition	Follow-up (mo)	N analyzed	Method, adjusted for factors* / unadjusted	Estimate (95% CI)†
Den Boer 2006[16]	No deficits versus sensory or motor deficits versus sensory and motor deficits (L4, L5, and S1 function assessed by physiotherapist)	Full return to work (100%) versus reduced work capacity (<100%; self-reported % of work capacity compared to before symptom onset)	6	182	I. Logistic regression adjusted for: fear of movement (TSK), passive pain coping (Pain Coping Inventory), negative outcome expectancies, physical workload, job satisfaction, duration of sick leave (wk), gender (male), age, education level (primary / secondary / tertiary), baseline disability (RMDQ) and pain (0–100 VAS), 3 d postoperative pain (0–100 VAS), analgesics week before surgery (types), duration (wk) and onset (acute) of symptoms; II. Logistic regression (unadjusted)	I. b = -.42, SE = .29, P = .12, OR = .64; II. b = -.51, P < .05; §OR = .60
	Symptom onset					
Den Boer 2006[16]	Acute versus nonacute onset of complaints	Full return to work (100%) versus reduced work capacity (<100%; self-reported % of work capacity compared to before symptom onset)	6	182	I. Logistic regression (unadjusted)	I. b = -.54, P > .05; §OR = .58
	Symptom duration					
Den Boer 2006[16]	Duration of the current episode of back and leg pain (wk)	Full return to work (100%) versus reduced work capacity (<100%; self-reported % of work capacity compared to before symptom onset)	6	182 (141 full, 41 reduced capacity)	I. Logistic regression (unadjusted)	I. b = -.01, P > .05; full M = 37.2, SD = 27.5, reduced M = 36.1, SD = 29.8; §OR = 1.07 (0.57–2.02)
Nguyen 2011[41]	Time from injury to surgery (d)	Return to work (part- or full-time) 2 y after surgery versus failure to return to work in any capacity	24	725	I. Logistic regression (unadjusted)	I. OR = .98 (0.97–1.00), P = .03
O'Donnell 2018[42]	Time from injury to surgery (mo)	Return to (same/different) work within 2 y sustained for ≥6 mo versus no sustained return to work	36	1286	I. Stepwise logistic regression adjusted for: opioid use before surgery (yes), legal representation (yes), psychiatric comorbidity (yes), mean household income ($)	I. OR = .98 (0.97–0.99), P < .01
Hagg 2003[23]	Duration of low back pain (y)	Work status at follow-up, working (part- or fulltime) versus not working	24	201 (74 working, 127 not working)	I. Independent t-test (unadjusted)	I. Working M = 8.6, SD = 7.5, not working M = 7.4, SD = 6.4, P > .05; §d = .18 (-0.11 to 0.46)
	Spinal pathology					
Nguyen 2011[41]	Herniated disc / radiculopathy / spondylolisthesis /spinal stenosis versus degenerative disc disease	Return to work (part- or full-time) 2 y after surgery versus failure to return to work in any capacity	24	725	I. Logistic regression (unadjusted)	I. P = .41; herniated disc OR = .81 (0.56–1.18), P = .27; radiculopathy OR = 2.00 (0.70–5.75), P = .20; spondylolisthesis OR = .80 (0.37–1.34), P = .58; spinal stenosis OR = 1.75 (0.11–28.45), P = .70
Anderson 2015[4]	Spondylosis versus no spondylosis	Return to work within 2 y after surgery sustained for ≥6 mo versus failure to return to work	36	2799 (199 spondylosis)	I. Logistic regression adjusted for: depression (yes), working within a week before surgery (yes), age (>50), opioid use for >1 y before surgery (yes), legal representation (yes), another lumbar surgery after index fusion	I. OR = .65 (0.43–0.96), P < .027

(*continued on next page*)

Table 2. Continued

Study ID	Prognostic factor	Outcome		Sample size	Analysis	Effect estimates
	Measure, definition	Measure, definition	Follow-up (mo)	N analyzed	Method, adjusted for factors* / unadjusted	Estimate (95% CI)†
					(yes), type of fusion (ALIF, PLF + PLIF), graft type (allograft)	
	Comorbidity					
Hagg 2003[23]	Comorbidity versus no comorbidity	Work status at follow-up, working (part- or fulltime) versus not working	24	196 (78 comorbidity, 118 no comorbidity)	I. 2 × 2 frequency table (Chi-squared test) (unadjusted)	I. 25 comorbidity, 48 no comorbidity were working at follow-up, $P > .05$; §OR = .68 (0.37–1.23)
	BMI					
Nguyen 2011[41]	BMI (kg/m^2)	Return to work (part- or full-time) 2 y after surgery versus failure to return to work in any capacity	24	725	I. Logistic regression (unadjusted)	I. OR = .99 (0.96–1.02), $P = .42$
	Smoking					
Anderson 2006[5]	Smoking versus not smoking	Work status at follow-up, working versus not working	24 (mean 30)	106	I. Logistic regression adjusted for: baseline work status (working), gender (NR), worker's compensation (yes), age <48, age >48, baseline pain (0–10 VAS), baseline disability (RMDQ), levels fused (single), cage type (BAK)	I. OR = .96 (0.31–3.02), $P = .94$
Hagg 2003[23]	Smoking versus not smoking	Work status at follow-up, working (part- or fulltime) versus not working	24	200 (84 smoking, 116 not smoking)	I. 2 × 2 frequency table (Chi-squared test) (unadjusted)	I. 32 smoking, 42 not smoking were working at follow-up, $P > .05$; §OR = 1.09 (0.61–1.95)
Nguyen 2011[41]	Current smoker / ex-smoker versus never a smoker	Return to work (part- or full-time) 2 y after surgery versus failure to return to work in any capacity	24	725	I. Logistic regression (unadjusted)	I. $P = .008$; current smoker OR = .53 (0.36–0.79), $P = .002$; ex-smoker OR = .70 (0.36–1.34), $P = .30$
	Analgesics use					
O'Donnell 2018[42]	No opioid use versus short-term (0–14 d) versus moderate (15–90 d) versus long-term opioid use (>90 d) before surgery	Return to (same/different) work within 2 y sustained for ≥6 mo versus no sustained return to work	36	1286 (566 no opioid, 126 short-term, 315 moderate, 279 long-term)	I. Analysis of variance, Bonferroni-corrected post hoc t-tests (unadjusted)	I. 64% (363) no opioid, 64% (80) short-term, 53% (166) moderate, 37% (103) long-term returned to work; effect of opioid group $P < .01$: no opioid > moderate $P < .01$, no opioid > long-term $P < .01$, short-term > long-term $P < .01$, moderate > long-term $P < .01$, no opioid versus short-term $P = 1.00$, short-term versus moderate $P = .21$
	No opioid use versus opioid use before surgery			1286 (566 no opioid, 720 opioid)	II. Stepwise logistic regression adjusted for: time from injury to surgery (mo), legal representation (yes), psychiatric comorbidity (yes), mean household income ($)	II. OR = .54 (0.39–0.75), $P < .01$
Anderson 2015[4]	Prescription opioid analgesia for >1 y before surgery versus none or <1 y	Return to work within 2 y after surgery sustained for ≥6 mo versus failure to return to work	36	2799 (637 opioids, 2162 no opioids)	I. Logistic regression adjusted for: depression (yes), working within a week before surgery (yes), age (>50), legal representation (yes), spinal pathology (spondylosis), another lumbar surgery after index fusion (yes), type of fusion (ALIF, PLF + PLIF), graft type (allograft)	I. OR = .58 (0.43–0.80), $P < .001$
Den Boer 2006[16]	Regular medications in the past week: no analgesics	Full return to work (100%) versus reduced work capacity	6	182	I. Logistic regression (unadjusted)	I. b = -.10, $P > .05$; §OR = .90

(continued on next page)

Table 2. Continued

Study ID	Prognostic factor	Outcome		Sample size	Analysis	Effect estimates
	Measure, definition	Measure, definition	Follow-up (*mo*)	N analyzed	Method, adjusted for factors* / unadjusted	Estimate (95% CI)†
	versus acetaminophen versus NSAID versus combination of NSAID and acetaminophen versus opioid versus combination of opioid and acetaminophen or NSAID	(<100%; self-reported % of work capacity compared to before symptom onset)				
	Psychological					
	Psychiatric comorbidity					
O'Donnell 2018[42]	Having versus not having any psychiatric comorbidity (depression, anxiety, adjustment disorder, bipolar disorder, PTSD, schizophrenia)	Return to (same/different) work within 2 y sustained for ≥6 mo versus no sustained return to work	36	1286 (37 with, 1249 without psychiatric comorbidity)	I. Stepwise logistic regression adjusted for: opioid use before surgery (yes), time from injury to surgery (mo), legal representation (yes), household income ($)	I. OR = .36 (0.14–0.90), *P* = .02
	Anxiety					
Schade 1999[49]	Anxiety subscale from Psychological general well-being index	Return to any work	24 (23–30)	42	I. Parametric univariate analysis (unadjusted)	I. *P* < .05
	Fear of movement					
Den Boer 2006[16]	TSK score (13–52)	Full return to work (100%) versus reduced work capacity (<100%; self-reported % of work capacity compared to before symptom onset)	6	182 (141 full, 41 reduced capacity)	I. Logistic regression adjusted for: passive pain coping (Pain Coping Inventory), negative outcome expectancies, physical workload, job satisfaction, duration of sick leave (wk), gender (male), age, education level (primary / secondary / tertiary), baseline disability (RMDQ) and pain (0–100 VAS), 3 d postoperative pain (0–100 VAS), analgesics week before surgery (types), neurologic deficits (none / sensory or motor, both), duration (wk) and onset (acute) of symptoms; II. Logistic regression (unadjusted)	I. b = -.09, SE = .04, *P* = .03, OR = .92; II. b = -.11, *P* < .01; full M = 39.2, SD = 6.3, reduced M = 42.1, SD = 5.8; §OR = .43 (0.23–0.81)
	Depression					
Schade 1999[49]	Depression subscale from Psychological general well-being index	Return to any work	24 (23–30)	42	I. Stepwise regression adjusted for: baseline pain/disability (NR), occupational mental stress (sum index); II. Parametric univariate analysis (unadjusted)	I. Beta = -.37, T = 2.63, *P* < .01; §OR = .69; II. *P* < .05
Hagg 2003[23]	Zhung Depression Scale (transformed from 20–80 to 0–100, higher scores indicating more severe depressive symptoms)	Work status at follow-up, working (part- or fulltime) versus not working	24	198 (72 working, 126 not working)	I. Independent t-test (unadjusted)	I. Working M = 37, SD = 13.9, not working M = 40, SD = 13, *P* > .05; §d = -.23 (-0.52 to 0.07)
Anderson 2015[4]	Clinically diagnosed depression (ICD-9 codes for depressive disorder, major depressive disorder,	Return to work within 2 y after surgery sustained for ≥6 mo versus failure to return to work	36	2799 (123 depression, 2676 no depression)	I. Logistic regression adjusted for: working within a week before surgery (yes), age (>50 y), opioids use for >1 y before surgery (yes), legal representation (yes), spinal pathology (spondylosis), another lumbar surgery after index fusion (yes), type of fusion (ALIF, PLF + PLIF),	I. OR = .38 (0.20–0.72), *P* < .002; II. 13 with depression, 884 without depression returned to work; §OR = .24 (0.13–0.43)

(*continued on next page*)

Table 2. Continued

Study ID	Prognostic factor	Outcome		Sample size	Analysis	Effect estimates
	Measure, definition	Measure, definition	Follow-up (mo)	N analyzed	Method, adjusted for factors* / unadjusted	Estimate (95% CI)†
	dysthymic disorder) versus no diagnosis of depression				graft type (allograft); II. 2 × 2 frequency table (unadjusted)	
	Vitality					
Schade 1999[49]	Vitality subscale from Psychological general well-being index	Return to any work	24 (23–30)	42	I. Parametric univariate analysis (unadjusted)	I. $P < .05$
	Coping					
Den Boer 2006[16]	Passive pain coping subscale score (21–84; Pain-Coping Inventory)	Full return to work (100%) versus reduced work capacity (<100%; self-reported % of work capacity compared to before symptom onset)	6	182 (141 full, 41 reduced capacity)	I. Logistic regression adjusted for: fear of movement (TSK), negative outcome expectancies, physical workload, job satisfaction, duration of sick leave (wk), gender (male), age, education level (primary / secondary / tertiary), baseline disability (RMDQ) and pain (0–100 VAS), 3 d postoperative pain (0–100 VAS), analgesics week before surgery (types), neurologic deficits (none / sensory or motor, both), duration (wk) and onset (acute) of symptoms; II. Logistic regression (unadjusted)	I. b = -.08, SE = .04, $P = .03$, OR = .93; II. b = -.10, $P < .001$; full M = 39.8, SD = 7.8, reduced M = 44.8, SD = 7.6; §OR = .31 (0.16–0.59)
	Expectancies					
Den Boer 2006[16]	Negative outcome expectancies score (5–20; 5 items scored on 4-point Likert scale)	Full return to work (100%) versus reduced work capacity (<100%; self-reported % of work capacity compared to before symptom onset)	6	182 (141 full, 41 reduced capacity)	I. Logistic regression (unadjusted)	I. b = -.18, $P > .05$; full M = 10.2, SD = 1.7, reduced M = 10.7, SD = 1.5; §OR = .58 (0.31–1.09)
	Personality					
Hagg 2003[23]	Neuroticism subscale from KSP (standardized T score)	Work status at follow-up, working (part- or fulltime) versus not working	24	189 (73 working, 115 not working)	I. Independent t-test (unadjusted)	I. Working M = 49.8, SD = 8.3, not working M = 52.9, SD = 9.3, $P = .02$; §d = -.35 (-0.64 to -0.05)
	Social introversion subscale from KSP (standardized T score)					I. Working M = 45.1, SD = 7.3, not working M = 46.4, SD = 7.4, $P > .05$; §d = -.18 (-0.47–0.12)
	Aggressiveness subscale from KSP (standardized T score)					I. Working M = 51.1, SD = 4.6, not working M = 50.3, SD = 4.5, $P > .05$; §d = .18 (-0.12 to 0.47)
	Impulsiveness subscale from KSP (standardized T score)					I. Working M = 50.1, SD = 8.9, not working M = 48.4, SD = 8.9, $P > .05$; §d = .19 (-0.10–0.49)
	Cluster A (paranoid, schizoid, schizotypal) personality disorder score from modified SCID-II (excluding antisocial personality)			140 (62 working, 78 not working)		I. Working M = 4.5, SD = 2.9, not working M = 4.2, SD = 3.1, $P > .05$; §d = .10 (-0.23 to 0.43)

(*continued on next page*)

Table 2. Continued

Study ID	Prognostic factor	Outcome		Sample size	Analysis	Effect estimates
	Measure, definition	Measure, definition	Follow-up (mo)	N analyzed	Method, adjusted for factors* / unadjusted	Estimate (95% CI)†
	Cluster B (borderline, histrionic, narcissistic) personality disorder score from modified SCID-II (excluding antisocial personality)					I. Working M = 5.0, SD = 3.6, not working M = 5.1, SD = 4.2, $P > .05$; §d = -.03 (-0.36 to 0.31)
	Cluster C (avoidant, dependent, obsessive-compulsive, passive-aggressive) personality disorder score from modified SCID-II (excluding antisocial personality)					I. Working M = 5.9, SD = 4.7, not working M = 6.1, SD = 4.69, $P > .05$; §d = -.04 (-0.38 to 0.29)
	Any personality disorder versus no personality disorder (SCID-II)			120 (1 personality disorder, 119 no personality disorder)	II. 2 × 2 frequency table (Fisher's exact test) (unadjusted)	II. 0 personality disorder, 54 no personality disorder were working at follow-up; §*not possible to calculate OR*
	Pain behavior					
Hagg 2003[23]	UAB Pain Behavior Scale score (0–10)	Work status at follow-up, working (part- or fulltime) versus not working	24	196 (74 working, 122 not working)	I. Independent t-test (unadjusted)	I. Working M = 2.2, SD = 1.4, not working M = 2.5, SD = 1.6, $P > .05$; §d = -.20 (-0.49 to 0.09)
	Waddell inappropriate signs and symptoms test score (0–10)			200 (74 working, 126 not working)		I. Working M = .8, SD = 1.3, not working M = .8, SD = 1.3, $P > .05$; §d = .00 (-0.29 to 0.29)
	Pain Drawing					
Hagg 2003[24]	Pain Drawing (Gatchel et al, 1986 scoring, 0–256)	Work status at follow-up, working (part- or fulltime) versus not working	24	201 (76 working, 125 not working)	I. Mann-Whitney U test (unadjusted)	I. Working M = 27, SD = 21.4, not working M = 25, SD = 20.2, $P > .05$; §d = .10 (-0.19 to 0.38)
	Pain Drawing (Sivik et al, 1992 scoring) nonorganic (high somatization tendency, ≥6/20) versus organic (low somatization tendency, <6/20)			201 (97 nonorganic, 104 organic)	II. Fisher's exact test (unadjusted)	II. 38 nonorganic, 38 organic were working at follow-up, $P = 1.0$, §OR = 1.12 (0.63–1.98)
	Pain Drawing (Ransford et al, 1976 Penalty Points scoring) nonorganic (abnormal, ≥3/15) versus organic (normal, <3/15)			201 (52 nonorganic, 149 organic)		II. 19 nonorganic, 37 organic were working at follow-up, $P = .9$, §OR = .93 (0.48–1.79)
	Pain Drawing (Udén et al, 1988 scoring) nonorganic (nonorganic, possibly nonorganic) versus			201 (77 nonorganic 124 organic)		II. 33 nonorganic, 43 organic were working at follow-up, $P = 1.0$, §OR = 1.41 (0.79–2.53)

(*continued on next page*)

Table 2. Continued

Study ID	Prognostic factor	Outcome		Sample size	Analysis	Effect estimates
	Measure, definition	Measure, definition	Follow-up (mo)	N analyzed	Method, adjusted for factors* / unadjusted	Estimate (95% CI)†
	organic (organic, possibly organic)					
	Job satisfaction					
Schade 1999[49]	4-item General job satisfaction scale (1–5 Likert ratings)	Return to any work	24 (23–30)	42	I. Parametric univariate analysis (unadjusted)	I. $P < .01$
Den Boer 2006[16]	Job satisfaction scale score (13–65)	Full return to work (100%) versus reduced work capacity (<100%; self-reported % of work capacity compared to before symptom onset)	6	182	I. Logistic regression adjusted for: fear of movement (TSK), passive pain coping (Pain Coping Inventory), negative outcome expectancies, physical workload, duration of sick leave (wk), gender (male), age, education level (primary / secondary / tertiary), baseline disability (RMDQ) and pain (0–100 VAS), 3 d postoperative pain (0–100 VAS), analgesics week before surgery (types), neurologic deficits (none / sensory or motor, both), duration (wk) and onset (acute) of symptoms; II. Logistic regression (unadjusted)	I. b = -.02, SE = .04, $P = .64$, OR = .98; II. b = .49, $P < .05$; §OR = 2.18 (1.15–4.11)
	Job-related resignation					
Schade 1999[49]	4-item Job-related resignation scale (1–5 Likert ratings)	Return to any work	24 (23–30)	42	I. Parametric univariate analysis (unadjusted)	I. $P < .05$
	Occupational stress					
Schade 1999[49]	Sum index of occupational mental stress (3 items rated none / mild / severe)	Return to any work	24 (23–30)	42	I. Stepwise multiple regression adjusted for: baseline pain/disability (NR), depression (subscale from Psychological general well-being index); II. Parametric univariate analysis (unadjusted)	I. Beta = -.28, T = 2.13, $P > .05$; §OR = .76; II. $P < .05$

Abbreviations: ALIF, anterior lumbar interbody fusion; b, unstandardized regression coefficient; BAK, Bagby and Kuslich cage; beta, standardized regression coefficient; BMI, body mass index; CI, confidence interval; ICD, International Classification of Diseases; KSP, Karolinska Scales of Personality; M, mean; NR, not reported; NSAID, nonsteroid anti-inflammatory drugs; ODI, Oswestry Disability Index; OR, odds ratio; PLF, posterior lumbar fusion; PLIF, posterior lumbar interbody fusion; RMDQ, Roland Morris Disability Questionnaire; SCID-II, Structured Clinical Interview for DSM-IV axis II disorders; SD, standard deviation; SE, standard error; TSK, Tampa Scale for Kinesiophobia; VAS, Visual Analog Scale.
*If categories not specified, factor analyzed as a continuous variable; nonsignificant PFs removed in a stepwise manner in adjusted analyses are not listed, but included: worker's compensation, being on sick leave, neuroticism;21 age, BMI, diagnoses, education, fusion approach, sex, marital status, N of levels with degenerative changes, N of rehabilitation / vocational sessions;39 age, gender, labor occupation, marital status, permanent disability, psychotherapy / physiotherapy / chiropractic care use;40 for Schade 1999, only the final model including significant PFs from domain-specific models is reported, but the following nonsignificant PFs were removed from within-domain stepwise models: demographics, low back pain history, physical and MRI findings (medical); anxiety, self-controls, wellbeing, general health, vitality (psychological factors); job satisfaction, social support (psychosocial aspects of work).[47]
†Statistics reported where available. 95% confidence interval provided in brackets where available. The sign of bs from Den Boer 2006 was reversed along with the outcome definition for consistency with the remaining studies. The sign of betas from Schade 1999 was also reversed as the context and narrative indicate that the reported factors were negative predictors of return to work and thus original prediction was likely made towards negative outcome.
‡Unclear if the PF was a composite measure of pain and disability or one of these factors, and which was adjusted for in the analysis.
§Effects estimated from available data where possible.

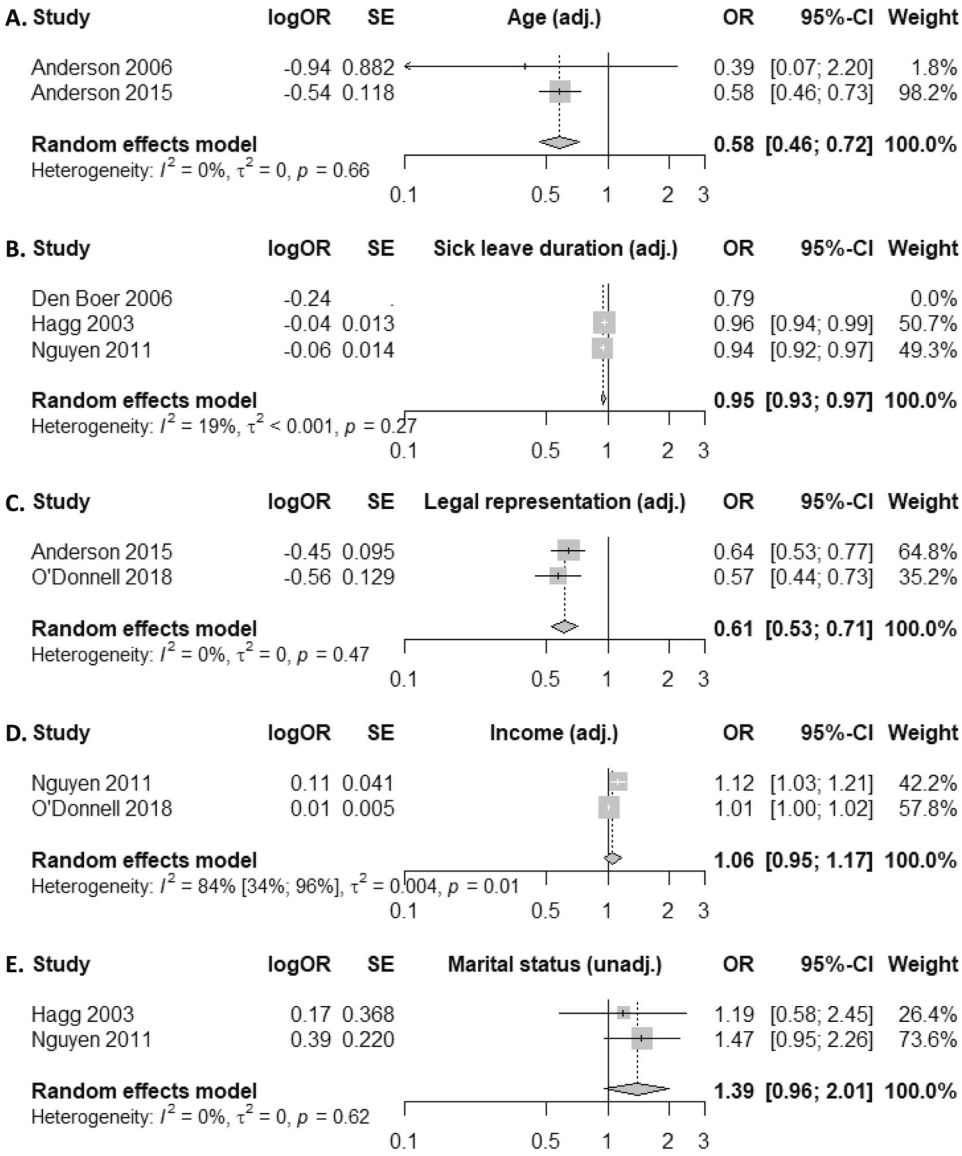

Figure 3. Forest plots of pooled effects of (A) age, (B) duration of sick leave, (C) legal representation, (D) income, and (E) marital status on return to work after surgery. If standard error or confidence interval of effect estimate was missing, that effect was not included in meta-analysis, but its magnitude is presented alongside other eligible studies (B). Abbreviations: adj., adjusted; CI, confidence interval; logOR, log-odds ratio; OR, odds ratio; SE, standard error; unadj., unadjusted effect.

showing a small significant negative relationship with work capacity in adjusted analysis.[16]

We found no significant independent associations between RTW and 2 work-related psychological factors, that is, *job satisfaction* (1 low RoB study)[16] and *occupational mental stress* (1 high RoB study).[49]

The prognostic value of several psychological factors was only examined in unadjusted analyses. A single high RoB study[49] reported significant negative relationships of *anxiety* and *job-related resignation*, and a positive relationship of *vitality,* with RTW after surgery, although effect estimates were not available. There were no significant associations between RTW and other psychological factors, including *pain behavior* (1 low RoB study),[23] *pain drawing* (1 high RoB study),[24] negative outcome *expectancies* (1 low RoB study),[16] and *personality* traits and disorders (1 low RoB study;[23] except for a small unadjusted negative effect of neuroticism).

Quality of Evidence

An overview of GRADE judgements of the quality of evidence is presented in Figure 4, and more detailed assessment is available in Supplementary Table S3.

There was *moderate*-quality evidence that demographic, socioeconomic, and affective psychological factors predict RTW. In particular, older age demonstrated independent negative prognostic value, gender, however, was unrelated to RTW outcomes. Regarding socioeconomic factors, participants who were employed before surgery were more likely to RTW even after accounting for potential confounders, whereas longer duration of sick leave, higher physical workload, and having legal representation showed independent negative associations with RTW. Income, workers' compensation, and general education level did not predict RTW after adjusting for other factors. Among affective

	Return to work		
Baseline predictors	*N of cohorts (patients)*	*Effect*	*Overall quality*
Sociodemographic			
Demographic	4 (3288)	✓	●●●○
Socioeconomic	5 (4574)	✓	●●●○
Health-related			
Pain	3 (487)	✗	●○○○
Disability	4 (529)	✗	●○○○
Spinal pathology	3 (3182)	✓	●○○○
Symptom duration	4 (2394)	✓	●●○○
Comorbidities	3 (1001)	✗	●○○○
Analgesics	3 (4267)	✓	●●○○
Psychological			
Affective	4 (4325)	✓	●●●○
Pain-related	2 (383)	✓	●○○○
Personality	1 (189)	✗	●○○○
Work-related	2 (224)	✗	●○○○

Sig. association:
Yes ✓, No ✗

Quality of evidence:
●●●● High
●●●○ Moderate
●●○○ Low
●○○○ Very low

Figure 4. Overall quality of evidence for the reviewed associations according to Grading of Recommendations, Assessment, Development and Evaluations (GRADE) framework.[22,34]

psychological factors, particularly depression and having any psychiatric comorbidity were found to be important independent negative predictors of RTW.

There was also *low*-quality evidence for independent prognostic value of symptom duration and analgesics use. Specifically, participants with longer duration of CLBP and those with opioids prescription and using opioids for longer were less likely to RTW after surgery.

The quality of evidence for the remaining associations was *very low*. Type of spinal pathology, in particular presence of spondylosis, and pain-related psychological factors such as passive coping and fear of movement, were identified as potential independent negative predictors of RTW. Finally, preoperative disability, pain intensity or pain in response to movement or touch, comorbidities, personality, and work-related psychological factors appeared unrelated to RTW outcomes.

The reasons for downgrading the quality of evidence can be summarized as follows. The data came largely from exploratory phase-1 studies, and only 3 of the examined relationships were supported by confirmatory phase-2 studies (analgesics use, depression, and work status).[4,5,42] Lack of adjustment for any confounders further affected the confidence in some examined associations, especially in the psychological domain. Another issue that decreased the quality of evidence and limited the opportunities for its quantitative synthesis was imprecision related to inadequate sample size and insufficient results reporting, affecting half of the reviewed relationships. Limited number of studies for several candidate predictors presented potential publication bias.

Discussion

We systematically reviewed the evidence for preoperative predictors of RTW after spinal surgery for CLBP, and performed narrative and quantitative synthesis of results where possible. Our main findings indicate that sociodemographic and affective psychological factors likely predict RTW. Symptom duration and opioid analgesic use also have potential prognostic value. The evidence for other health-related and psychological predictors is less certain.

Sociodemographic Predictors

We found moderate-quality evidence supported by meta-analysis that older age independently predicts decreased likelihood of RTW after spinal surgery, in line with previous systematic reviews.[35,43] This seems particularly relevant in light of increasing prevalence of CLBP with age. Its burden will likely increase with aging population,[32] and so may the number of people in need of spinal surgery.

The independent prognostic value of several socioeconomic work-related factors is supported by moderate-quality evidence. Patients working before surgery, with shorter duration of sick leave, lower physical workload, and without legal representation are more likely to RTW after surgery. These findings, with the effects of sick leave and legal representation confirmed in meta-analyses, reinforce the consistent evidence from previous reviews regarding RTW after surgery based largely on unadjusted associations,[35,43] and from noninterventional studies on RTW in CLBP.[50] We further found that income may be less relevant to RTW after adjusting for stronger predictors such as legal representation or opioids use (although this evidence may be limited by statistical heterogeneity), and that the previously reported effect of workers' compensation claims on RTW[35] may lose its predictive ability after adjusting for work status. The value of socioeconomic predictors of RTW may depend on national healthcare and insurance systems, yet, the limited number of studies prevented sensitivity analysis to explore this further. Effects of legal representation, income, and work status were only assessed in US-based studies, however, association between sick leave and RTW was found across US, Swedish, and Dutch studies.

Health-Related Predictors

Pertinent to our review question regarding CLBP, we found low-quality evidence for an independent effect of longer symptom duration on reduced RTW, further supported by their pooled unadjusted association. These findings are in agreement with a previous review of RTW rates after spinal surgery,[35] and with strong evidence from nonsurgical CLBP studies that delay in referral for intervention has adverse effects on RTW.[50] Preoperative opioid prescription and prolonged opioid use also independently predicted reduced RTW in the current review (low-quality evidence). Opioids are commonly prescribed for moderate and severe pain that could not be managed with other treatments, therefore these patients likely represent more severe cases.[19] Negative side effects of prolonged opioid use could further interfere with RTW.[17] Very low-quality evidence suggests that spinal pathology, particularly presence of spondylosis, may be a negative independent predictor of RTW. Patients with spondylosis are often older and present with more persistent symptoms. Overall, these prognostic effects, in line with the duration of sick leave, appear to suggest that chances of RTW decrease with increasing duration and severity of CLBP. However, preoperative disability (seemingly closely related to functional recovery indexed by RTW) only predicted this outcome in unadjusted meta-analysis, but not after adjusting for socioeconomic and psychological factors (very low-quality evidence).

Psychological Predictors

Recommendations regarding the prognostic value of psychological distress for spinal surgery outcomes are scarce and vary across countries. For instance, in the case of disc herniation with radiculopathy, the UK National Institute for Health and Care Excellence advises not to use such information during patient selection for surgery, whereas the North American Spine Society suggests that patients should be assessed for signs of psychological distress, such as somatization and depression, based on fair evidence that these signs predict worse outcomes.[15,37] Our review, considering a broader range of degenerative spine diseases, found no evidence for pain behavior or nonorganic signs, but passive pain coping and fear of movement may be related to RTW after surgery (very low-quality evidence). Importantly, we found moderate-quality evidence that patients who are depressed or have a psychiatric comorbidity are less likely to RTW, even after controlling for potential confounders. Although the data was not suitable for meta-analysis, this evidence was supported by moderate effect sizes from large studies, including a confirmatory investigation. Unadjusted effects of anxiety and low vitality were consistent with the conclusion that negative affective factors likely predict reduced RTW.

Strengths and Limitations

This review provides a comprehensive evaluation of 33 candidate predictors for RTW, objectively reflecting functional recovery after surgery. We expand upon the previous literature[35,43] by including a broader range of degenerative diseases of the lumbar spine, while specifically focusing on patients with chronic pain. Identifying a larger set of relevant studies allowed to pool the effects of some associations in meta-analysis. Furthermore, our robust quality assessments provide a transparent overview of the certainty in available prognostic evidence in this filed, highlight prevalent methodological issues, and signpost directions for further research.

Low quality of evidence was a concern for predictors other than sociodemographic and affective psychological factors. Specific shortcomings of the available evidence are outlined in the *Risk of bias* and *Quality of evidence* results sections. The current review also has some limitations. First, several examined associations included indirect evidence, where it was not possible to verify chronic pain status of all participants,[4,16,42] thus these samples may not accurately reflect the review question. Nonetheless, available data strongly suggested their eligibility and any uncertainty was reflected in the quality assessments. Second, in 2 included studies, not all patients were working before surgery.[5,23,24] While this may seem suboptimal considering RTW outcome, it allowed us to examine the prognostic effect of preoperative work status. Third, definitions of RTW were not consistent across all studies, with some specifying sustained RTW, and others work capacity. Although all these definitions were relevant to our review question, they added a degree of study heterogeneity. However, outcome definition was not identified as a potential source of inconsistencies in the results. Fourth, the reviewed evidence comes from various American

and European economic contexts, adding heterogeneity of the healthcare and compensation systems. Ideally, the impact of the above-mentioned limitations should be examined in sensitivity analyses, yet these were not possible due to the small number of studies. This also limits our confidence in the precision of pooled estimates, based only on 2 to 3 studies for each association examined in meta-analysis. It also affected the feasibility of quantitative synthesis for other candidate predictors. Moreover, moderate or large effect sizes contributed only to prognostic effects of socioeconomic and affective factors, while the magnitude of all pooled estimates was small. Finally, the current article is part of a broader review including predictors of patient-reported pain and disability outcomes. While the decision to report them in a separate article was not made a priori, it allows to provide more in-depth assessment of these distinct outcomes.

Implications

The identified limitations of the existing evidence suggest specific directions for further research. Issues preventing meta-analysis could be addressed at the levels of study design and results reporting, where consensus on consistent measures and definitions of the same predictors and outcomes would reduce heterogeneity[11,20]; monitoring and reporting attrition could ensure representativeness of the studied samples; and transparent reporting of both positive and negative results with precision estimates would allow quantitative synthesis and increase certainty in presented evidence. This review further highlights the importance of controlling for potential alternative explanations in prognosis research. For instance, possible effects of education level, workers' compensation, disability, and work-related psychological factors were found in unadjusted analyses, but could not be replicated after adjusting for other factors. As there is no recommended set of relevant confounders, we suggest that age, socioeconomic, and affective factors should be adjusted for in future prognosis research. Additional high-quality confirmatory studies should verify the independent prognostic value of symptom duration, analgesics use, pain-related psychological factors, and spinal pathology, which is currently supported by low or very low-quality evidence. Despite overall small effect sizes arising from the current syntheses, combining several prognostic factors could increase the accuracy of outcome prediction, thus the identified likely predictors of RTW after spinal surgery should be considered in the development of clinical prediction models.

There are several practical implications of the presented findings. For instance, the negative effects of socioeconomic factors and symptom duration on RTW suggest that patients might benefit from being operated earlier in the course of CLBP, or working as long as possible until surgery but perhaps with reduced workload. Furthermore, patients with CLBP with signs of depression or maladaptive pain coping may benefit from preoperative cognitive-behavioral therapy.[48] Ideally, multidisciplinary approaches to pain management should be sought as an alternative to opioid pain relief, which, if necessary, should only be prescribed for short periods.

Conclusions

The likelihood of RTW after spinal surgery for CLBP appears to depend on patients' demographics, socioeconomic situation, medical history, and affective psychological characteristics. We found likely negative prognostic value of older age, longer sick leave, having legal representation, higher physical workload, not working before surgery, having psychiatric comorbidities and depression for RTW. Longer symptom duration and opioid use also potentially predict reduced RTW, whereas the prognostic value of other preoperative factors is less certain. The current level of evidence may not be sufficient for the development of clinical guidelines regarding prognosis, and more high-quality prospective data would increase confidence in the above associations.[56] However, the identified predictors could inform the design of future confirmatory studies and clinical prediction models, help to estimate the likelihood of functional recovery and choose the best course of treatment at the right time, to reach individual patient goals and maximize the benefit from surgery.

Acknowledgments

We thank Dr. Sarah Nevitt for her advice on the systematic review protocol.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.jpain.2022.02.003.

References

1. Altman DG: Practical Statistics for Medical Research. CRC Press; 1990

2. Anderson JT, Haas AR, Percy R, Woods ST, Ahn UM, Ahn NU: Clinical depression is a strong predictor of poor lumbar fusion outcomes among workers' compensation subjects. Spine 40:748-756, 2015

3. Anderson PA, Schwaegler PE, Cizek D, Leverson G: Work status as a predictor of surgical outcome of discogenic low back pain. Spine 31:2510-2515, 2006

4. Bakbergenuly I, Hoaglin DC, Kulinskaya E: Methods for estimating between-study variance and overall effect in meta-analysis of odds ratios. Res Synth Methods 11:426-442, 2020. Wiley Online Library

5. Balduzzi S, Rücker G, Schwarzer G: How to perform a meta-analysis with R: A practical tutorial. Evid Based

Ment Health 22:153-160, 2019. Royal College of Psychiatrists

6. Bramer WM: Reference checking for systematic reviews using Endnote. J Med Libr Assoc JMLA 106:542., 2018. Medical Library Association

7. Bruls VEJ, Bastiaenen CHG, de Bie RA: Prognostic factors of complaints of arm, neck, and/or shoulder: A systematic review of prospective cohort studies. Pain 156:765-788, 2015. 2015/02/07 ed.

8. Centre for Reviews and Dissemination: Systematic Reviews: CRD's Guidance for Undertaking Reviews in Healthcare. University of York: CRD; 2009

9. Chiarotto A, Boers M, Deyo RA, Buchbinder R, Corbin TP, Costa LO, Foster NE, Grotle M, Koes BW, Kovacs FM: Core outcome measurement instruments for clinical trials in nonspecific low back pain. Pain 159:481., 2018. Wolters Kluwer Health

10. Cohen J: A coefficient of agreement for nominal scales. Educ Psychol Meas 20:37-46, 1960. Sage Publications Sage CA: Thousand Oaks, CA

11. Costa L da CM, Maher CG, Hancock MJ, McAuley JH, Herbert RD, Costa LOP: The prognosis of acute and persistent low-back pain: A meta-analysis. CMAJ 184:E613-E624, 2012. CMAJ

12. Dagenais S, Caro J, Haldeman S: A systematic review of low back pain cost of illness studies in the United States and internationally. Spine J 8:8-20, 2008

13. de Campos TF: Low back pain and sciatica in over 16s: Assessment and management NICE Guideline [NG59]. J Physiother 63:120, 2017

14. den Boer JJ, Oostendorp RA, Beems T, Munneke M, Evers AW: Reduced work capacity after lumbar disc surgery: The role of cognitive-behavioral and work-related risk factors. Pain 126:72-78, 2006. 2006/07/18 ed

15. Deyo RA, Korff MV, Duhrkoop D: Opioids for low back pain. BMJ 350:g6380., 2015. British Medical Journal Publishing Group

16. Dhondt E, Van Oosterwijck J, Cagnie B, Adnan R, Schouppe S, Van Akeleyen J, Logghe T, Danneels L: Predicting treatment adherence and outcome to outpatient multimodal rehabilitation in chronic low back pain. J Back Musculoskelet Rehabil 33:277-293, 2020. 2019/07/30 ed.

17. Effect size calculator [Internet]. [Accessed February 23, 2021]. Available at: https://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-Home.php

18. Dowell D, Haegerich TM, Chou R: CDC guideline for prescribing opioids for chronic pain—United States, 2016. JAMA 315:1624-1645, 2016

19. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N: Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 113:9-19, 2005

20. Grooten WJA, Tseli E, Ang BO, Boersma K, Stalnacke BM, Gerdle B, Enthoven P: Elaborating on the assessment of the risk of bias in prognostic studies in pain rehabilitation using QUIPS-aspects of interrater agreement. Diagn Progn Res 3:5., 2019. 2019/05/17 ed.

21. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A: GRADE guidelines: A new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol 64:380-382, 2011. Elsevier

22. Hägg O, Fritzell P, Ekselius L, Nordwall A: Predictors of outcome in fusion surgery for chronic low back pain. A report from the Swedish Lumbar Spine Study. Eur Spine J 12:22-33, 2003

23. Hägg O, Fritzell P, Hedlund R, Möller H, Ekselius L, Nordwall A: Pain-drawing does not predict the outcome of fusion surgery for chronic low-back pain: A report from the Swedish Lumbar Spine Study. Eur Spine J 12:2-11, 2003

24. Harrer M, Cuijpers P, Furukawa TA, Ebert DD: Doing Meta-Analysis With R: A Hands-On Guide[Internet]. 1st ed. Boca Raton, FL and London: Chapman & Hall/CRC Press; 2021. Available at: https://www.routledge.com/Doing-Meta-Analysis-with-R-A-Hands-On-Guide/Harrer-Cuijpers-Furukawa-Ebert/p/book/9780367610074. Accessed August 19, 2021.

25. Hayden JA, Côté P, Bombardier C: Evaluation of the quality of prognosis studies in systematic reviews. Ann Intern Med 144:427-437, 2006. American College of Physicians

26. Hayden JA, Dunn KM, van der Windt DA, Shaw WS: What is the prognosis of back pain? Best Pract Res Clin Rheumatol 24:167-179, 2010

27. Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C: Assessing bias in studies of prognostic factors. Ann Intern Med 158:280-286, 2013. American College of Physicians

28. Hayden JA, Wilson MN, Riley RD, Iles R, Pincus T, Ogilvie R: Individual recovery expectations and prognosis of outcomes in non-specific low back pain: Prognostic factor review. Cochrane Database Syst Rev 2019, 2019. [Internet]. [Accessed February 23, 2021]. Available at: https://www.ncbi.nlm.nih.gov/pubmed/31765487

29. Hegarty D, Shorten G: Multivariate prognostic modeling of persistent pain following lumbar discectomy. Pain Physician 15:421-434, 2012

30. Higgins JP, Thompson SG: Quantifying heterogeneity in a meta-analysis. Stat Med 21:1539-1558, 2002. Wiley Online Library

31. Hoy D: The global burden of low back pain: Estimates from the Global Burden of Disease 2010 Study. Ann Rheum Dis 73:968-974, 2014

32. Hoy D, Bain C, Williams G, March L, Brooks P, Blyth F, Woolf A, Vos T, Buchbinder R: A systematic review of the global prevalence of low back pain. Arthritis Rheum 64:2028-2037, 2012

33. Huguet A, Hayden JA, Stinson J, McGrath PJ, Chambers CT, Tougas ME, Wozney L: Judging the quality of evidence in reviews of prognostic factor research: Adapting the GRADE framework. Syst Rev 2:1-12, 2013. Springer

34. Huysmans E, Goudman L, Van Belleghem G, De Jaeger M, Moens M, Nijs J, Ickmans K, Buyl R, Vanroelen C, Putman K: Return to work following surgery for lumbar radiculopathy: A systematic review. Spine J 18:1694-1714, 2018. 2018/05/26 ed.

35. International Association for the Study of Pain: IASP Terminology [Internet]. 2017 [Accessed February 5, 2021].

Available at: https://www.iasp-pain.org/Education/Content.aspx?ItemNumber=1698#Centralsensitization

36. Kreiner DS, Hwang S, Easa J, Resnick DK, Baisden J, Bess S, Cho CH, DePalma MJ, Dougherty P, Fernand R, Ghiselli G, Hanna AS, Lamer T, Lisi AJ, Mazanec DJ, Meagher RJ, Nucci RC, Patel RD, Sembrano JN, Sharma AK, Summers JT, Taleghani CK, Tontz WL, Toton JF: Diagnosis and Treatment of Lumbar Disc Herniation With Radiculopathy. North American Spine Society 1-100, 2012

37. Lenhard W, Lenhard A: Computation of effect sizes [Internet]. 2017 [Accessed July 13, 2021]. Available at: http://rgdoi.net/10.13140/RG.2.2.17823.92329

38. Lüdecke D: esc: Effect size computation for meta analysis [Internet]. Zenodo; 2018 [Accessed August 22, 2021]. Available at: https://zenodo.org/record/1249218

39. Modic MT, Pavlicek W, Weinstein MA, Boumphrey F, Ngo F, Hardy R, Duchesneau PM: Magnetic resonance imaging of intervertebral disk disease. Clinical and pulse sequence considerations. Radiology 152:103-111, 1984

40. Nguyen TH, Randolph DC, Talmage J, Succop P, Travis R: Long-term outcomes of lumbar fusion among workers' compensation subjects: A historical cohort study. Spine 36:320-331, 2011

41. OCEBM levels of evidence — Centre for Evidence-Based Medicine (CEBM), University of Oxford [Internet]. [Accessed February 23, 2021]. Available at: https://www.cebm.ox.ac.uk/resources/levels-of-evidence/ocebm-levels-of-evidence

42. O'Donnell JA, Anderson JT, Haas AR, Percy R, Woods ST, Ahn UM, Ahn NU: Preoperative opioid use is a predictor of poor return to work in workers' compensation patients after lumbar diskectomy. Spine 43:594-602, 2018. 03622436

43. Oosterhuis T, Smaardijk VR, Kuijer PPF, Langendam MW, Frings-Dresen MHW, Hoving JL: Systematic review of prognostic factors for work participation in patients with sciatica. Occup Env Med 76:772-779, 2019. 2019/07/13 ed.

44. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D: The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ 372:n71., 2021. British Medical Journal Publishing Group

45. Rajaee SS, Bae HW, Kanim LEA, Delamarter RB: Spinal fusion in the United States: Analysis of trends from 1998 to 2008. Spine 37:67-76, 2012

46. Rice AS, Smith BH, Blyth FM: Pain and the global burden of disease. Pain 157:791-796, 2016. LWW

47. Riley RD, Moons KGM, Snell KIE, Ensor J, Hooft L, Altman DG, Hayden J, Collins GS, Debray TPA: A guide to systematic review and meta-analysis of prognostic factor studies. BMJ 364:k4597., 2019. 2019/02/01 ed.

48. Rolving N, Sogaard R, Nielsen CV, Christensen FB, Bünger C, Oestergaard LG: Preoperative cognitive-behavioral patient education versus standard care for lumbar spinal fusion patients: Economic evaluation alongside a randomized controlled trial. Spine 41:18-25, 2016

49. Schade V, Semmer N, Main CJ, Hora J, Boos N: The impact of clinical, morphological, psychosocial and work-related factors on the outcome of lumbar discectomy. Pain 80:239-249, 1999

50. Steenstra IA, Munhall C, Irvin E, Oranye N, Passmore S, Van Eerd D, Mahood Q, Hogg-Johnson S: Systematic review of prognostic factors for return to work in workers with sub acute and chronic low back pain. J Occup Rehabil 27:369-381, 2017

51. Van Der Windt DA, Simons E, Riphagen II, Ammendolia C, Verhagen AP, Laslett M, Devillé W, Deyo RA, Bouter LM, de Vet HC: Physical examination for lumbar radiculopathy due to disc herniation in patients with low-back pain. Cochrane Database Syst Rev 17:1-94, 2010. John Wiley & Sons, Ltd

52. Vasseljen O, Woodhouse A, Bjrngaard JH, Leivseth L: Natural course of acute neck and low back pain in the general population: The HUNT study. Pain 154:1237-1244, 2013

53. Veroniki AA, Jackson D, Viechtbauer W, Bender R, Bowden J, Knapp G, Kuss O, Higgins JP, Langan D, Salanti G: Methods to estimate the between-study variance and its uncertainty in meta-analysis. Res Synth Methods 7:55-79, 2016. Wiley Online Library

54. Weir S, Samnaliev M, Kuo T-C, Choitir CN, Tierney TS, Cumming D, Bruce J, Manca A, Taylor RS, Eldabe S: The incidence and healthcare costs of persistent postoperative pain following lumbar spine surgery in the UK: A cohort study using the Clinical Practice Research Datalink (CPRD) and Hospital Episode Statistics (HES). BMJ Open 7:e017585, 2017

55. Wu A, March L, Zheng X, Huang J, Wang X, Zhao J, Blyth FM, Smith E, Buchbinder R, Hoy D: Global low back pain prevalence and years lived with disability from 1990 to 2017: Estimates from the Global Burden of Disease Study 2017. Ann Transl Med 8:299, 2020

56. Wupperman R, Davis R, Obremskey WT: Level of evidence in Spine compared to other orthopedic journals. Spine 32:388-393, 2007. LWW

57. Zweig T, Enke J, Mannion AF, Sobottke R, Melloh M, Freeman BJ, Aghayev E, Spine Tango C: Is the duration of pre-operative conservative treatment associated with the clinical outcome following surgical decompression for lumbar spinal stenosis? A study based on the Spine Tango Registry. Eur Spine J 26:488-500, 2017. 2016/12/17 ed.

58. Zweig T, Hemmeler C, Aghayev E, Melloh M, Etter C, Röder C: Influence of preoperative nucleus pulposus status and radiculopathy on outcomes in mono-segmental lumbar total disc replacement: Results from a nationwide registry. BMC Musculoskelet Disord 12:275, 2011

59. Halicka M, Duarte R, Catherall S, Maden M, Coetsee M, Wilby M: Predictors of Pain and Disability Outcomes Following Spinal Surgery for Chronic Low Back and Radicular Pain. Clin J Pain, 2022. https://doi.org/10.1097/AJP.0000000000001033