REVIEW

Anne F. Mannion
Achim Elfering

Predictors of surgical outcome and their assessment

A. F. Mannion (✉)
Spine Unit, Schulthess Klinik,
Lengghalde 2, 8008 Zürich, Switzerland
E-mail: anne.mannion@kws.ch
Tel.: +41-44-3857584
Fax: +41-44-3857590

A. F. Mannion
Department of Rheumatology and
Institute of Physical Medicine,
University Hospital, Zürich, Switzerland

A. Elfering
Department of Psychology,
University of Berne, Berne, Switzerland

Abstract The relatively high rate of failed back surgery has prompted the search for "risk factors" to predict the result of spinal surgery in a given individual. However, the literature reveals few unequivocal predictors and they often explain a relatively low proportion of variance in outcome. This suggests that we have a long way to go before being able to rest easily, having refused someone surgery on the basis of unfavourable baseline characteristics. The best recommendation is to ensure, firstly, that the indication for surgery is absolutely clear-cut (i.e. that surgically remediable pathology exists) and then to consider the various factors that may influence the "typical" outcome. Consistent risk factors for a poor outcome regarding return-to-work include long-term sick leave/receipt of disability benefit. Hence, every effort should be made to keep the individual in the workforce, despite the ongoing symptoms and plans for surgery. In patients with a particularly heavy job, consultation with occupational physicians might later ease the patient's way back into the workplace. Patients with degenerative disorders and/or comorbidity should be counselled that few of them will have complete/lasting pain relief or a complete return to pre-morbid function. Patients with a high level of distress may benefit from psychological treatment, before and/or accompanying the surgical treatment. The opportunity (time), encouragement (education and positive messages), and resources (referral to appropriate support services) to modify risk factors that are indeed modifiable should be offered, and realistic expectations should be discussed with the patient before the decision to operate is made.

Keywords Predictors · Risk factors · Surgery · Outcome · Questionnaires

Introduction

Failed back surgery is a problem that has become sufficiently widespread to even warrant its own special conferences, with recent reviews reporting failure rates ranging from 5 to 50% [18]. The substantial suffering of patients with failed back surgery syndrome, the associated costs to society, and the not inconsiderable complication rates associated with spinal surgery, per se, have prompted the search for predictors of outcome, in an attempt to better identify individuals who are likely to benefit from surgery. The development of "pre-screening" tools has also been encouraged, to assist with the patient selection procedure and the promotion of realistic expectations on behalf of the patient [57, 65].

Over the last 10–15 years, numerous studies have sought to identify predictors of surgical outcome. However, before the results of these studies can be reviewed and summarized, it is important that the reader be made aware of the numerous factors that may influence the (at times discrepant) findings from these studies. These factors include:

– the design of the study and the statistical methods used to identify predictors;
– the outcome measures employed, the means by which a "successful outcome" is defined, and the proportion of patients in the investigated group that typically achieve a successful outcome;
– the number and type of predictor factors subjected to examination in any given study, and their prevalence within the group under investigation;
– the specific pathology or surgical procedure under investigation and the defining characteristics of the patients with that pathology.

These issues must be considered carefully, in order that the reader may appreciate the somewhat complicated nature of the topic and may develop the critical thinking required to interpret the results of the existing and future studies of predictors.

Predictor study designs

Retrospective studies

Retrospective studies ("looking back in time") involve the examination of results at a given time after surgery, analysed in relation to the available data recorded at baseline in the patients' medical records. Occasionally, additional data will be collected at the time of the investigation regarding the current status or the situation (recalled) as it was before the surgery. As this type of study is not pre-planned, the most appropriate baseline information has often not been obtained. Further, the method of data collection has not been standardized and data are frequently missing. Often, if some particular attribute has not been noted in the medical records, it is assumed not to have been present, e.g. [29], which is not always the case. It is also not known whether any putative risk factors may already have been used in clinical practice as part of the selection criteria, hence reducing both the frequency of observation of these factors and their chances of achieving a statistically significant association with the outcome. Finally, if the "predictors" themselves are collected retrospectively, then the post-surgical status can heavily bias the accuracy of recall of the situation before surgery. Retrospective studies are hence the least robust type of study for identifying risk factors.

Prospective studies

As their name suggests, prospective studies ("looking forward in time") involve characterizing the baseline status of patients before surgery, and then following up the patients over time, in order to relate the baseline characteristics to the ultimate outcome. Prospective studies are carried out with a planned research design, allowing sufficient information to be gathered to investigate the phenomenon in question. Patient record keeping may be modified, specific questionnaires may be introduced, or additional measurements may be taken during routine patient management in order to collect the necessary data. As patients are made aware that they are participating in a study, their permission can be sought to obtain additional contact information, in order to reduce losses to follow-up. The duration of the follow-up may influence both the proportion of patients showing a successful outcome and also the predictors of the outcome. With new surgical techniques/implants, a follow-up of at least 2 years is typically required, as the development of inadvertent or unexpected complications must be included as part of the overall outcome assessment. However, for other procedures, in which the operation seeks to remedy a physical, mechanical obstruction (e.g. decompression surgery), the outcome can arguably be examined as soon as the patient has recovered from the direct effects of the operation per se. This can be expected to occur within a maximum of 6 months after surgery. In fact, longer follow-ups in these patients, although of interest for documenting the natural history of treatment of the condition, may introduce additional sources of error as far as the identification of predictors is concerned. It is conceivable that other factors (independent of the original surgical intervention) may influence the patient's rating of "outcome" in the long term, especially if the latter is based on self-ratings of current pain, disability, or quality of life (see later; outcome measures). Further, it cannot be assumed that all the predictors assessed at the baseline are necessarily intrinsic to the patient and will remain stable over time; especially changes in work-status/social situation or the development of comorbidities may influence the patient's rating of their status at the time of the follow-up, such that the assessments made at baseline no longer accurately characterize the patient. Arguably, predictor analyses should be carried out at repeated follow-ups (to identify "stable" and consistent predictors), and any potentially important but "labile" variables should be re-assessed simultaneously.

Statistical methods used to identify predictors

The simplest, but also most limited, studies of predictors involve splitting patients into (usually) two groups—good outcome and bad outcome—and examining the significance of the difference between these groups in relation to the mean values or presence/absence of various baseline variables (sometimes referred to as "case-control" studies) for each putative predic-

tor. In doing so, one can learn something about the distinguishing features of the two groups, or the relative risk of a poor outcome in the presence of a given risk factor, but only little about an individual's overall likelihood of a certain outcome. Nonetheless, the variables identified may be of use in guiding the development of more precise screening tools in further studies. When the outcome is measured as a continuous variable (e.g. improvement in pain, reduction in disability score), correlational analyses may be performed between the individual putative predictors and the chosen outcome measure to identify the strength of the relationship between the two. With this simple bivariate approach, any interactions between the effects of two or more variables on the outcome will not be identified; sometimes, the combination of two variables may represent a risk factor, although neither alone is significant [80]. Further, carrying out repeated individual analyses tends to result in the identification of numerous risk factors that are, in reality, delivering the same information. For example, duration of symptoms and depression may each be identified as important predictors of outcome, but they are so closely related to one another—i.e. patients who have endured persisting symptoms are often those that become depressed—that they end up delivering almost identical information in explaining the variance in outcome. Only multivariate analyses can account for this overlap of information (sometimes referred to as "collinearity"), since variables are entered into the predictive model only if they are able to uniquely explain any variance in outcome, i.e. able to add something *more* to the predictive model.

The most commonly utilized multivariate analysis method is logistic multiple regression. This is used to predict categorical outcome (good/bad) from a range of continuous and categorical variables. Often, bivariate analyses (outcome vs each predictor) are carried out first, in order to determine which of the many predictor variables should be examined in the multivariate model (i.e. to reduce the number of variables entered). The aim of logistic regression analysis is to find a subset of explanatory variables that can be combined to best predict outcome. For future patients, values for these explanatory variables can then be entered into the logistic regression equation (or "predictive model") to estimate the likelihood of a good (or bad) outcome for the given patient (a practical example of this is given in [40]). The coefficients from the predictive model are sometimes used to assign the variables a "cut-off score" in order to develop simple screening tools for use in clinical practice. Clearly, the sensitivity, specificity, and positive and negative predictive value of such tools must be adequate, and the model must be validated in an independent cohort, or its reliability and generalizability to other clinical settings may be limited. This is rarely done. In interpreting the results of logistic regression

analyses, it is important to note how much variance in outcome is explained by the chosen predictor variables, rather than just whether the association achieves a given level of statistical significance. There are abundant studies in the literature in which highly statistically significant predictors of outcome have been identified, but on closer examination these actually explain a very low proportion of the variance in outcome (e.g. [54], 75, 76]). This makes such predictors of limited use in clinical practice.

What constitutes a "successful outcome?"

The proportion of patients that can be considered a success after surgery as well as the factors that might predict a good outcome depend to a large extent on how success is defined [3, 80]. The success of outcome is likely best considered in relation to the predominant aim of the surgery. Hence, for decompression surgery for a herniated disc or spinal stenosis, the most important outcome may be the reduction of leg pain or sensory disturbances and/or walking capacity, whereas for "chronic degenerative low back pain," the relief of low back pain (LBP) will primarily govern the degree of success. For all of these conditions, the ability to regain normal function in activities of daily living will also be of importance, although this typically follows with time once the main symptoms have resolved. In the case of deformity surgery, pain or disability may not be an issue, and factors other than symptoms (such as cosmetic appearance, prevention of progressive worsening, and associated systemic complications) may determine the "success" of surgery. The success may also depend on the age-group and working status of the group under investigation as well as the answer to the question "who's asking?"—when viewed from the economic point of view, outcomes concerned with work capacity may be of greatest importance for younger patients of working-age.

Global assessments' scores often give the most direct answer to the question "did the operation help?" and allow the patient to interpret the question in relation to his or her own particular pre-surgical problems and expectations of surgery. For the purposes of predictor studies, multiple response categories for this question (commonly between three and seven responses, ranging from "the surgery helped a lot" through to "the surgery made things worse," or "excellent result" through to "bad result") are often collapsed to dichotomize the data into "good" and "poor" outcome groups. Some authors consider that all responses greater than a "neutral" outcome (i.e. no change) should be considered as a positive result, while others argue that for elective surgical procedures a notable improvement should be required (i.e. more than "helped a little" or "fair result") to consider the operation a success [38].

In predictor studies in which continuous variables such as the Roland Morris score, Oswestry Disability Index, or pain visual analogue scales (VAS) are used as the primary outcome measure, some indication of the cut-off value corresponding to a ''good outcome'' is required, i.e. the value of the minimal clinically relevant change score. To determine the value of such cut-off scores, the method of Receiver Operating Characteristics (ROC) is commonly used. The ROC curve synthesizes information on sensitivity and specificity for detecting improvement (according to some dichotomized external criterion) for each of several possible cut-off points in change score [23]. Thus, sensitivity and specificity can be calculated for a change score of one point, two points, and so on. This method is analogous to evaluating the predictive power of a diagnostic test in which the instrument (questionnaire) change score is the diagnostic test and the global outcome (dichotomized as described above) is used to represent the gold standard [23]. Using such methods, it has been shown that a ''good outcome'' cut-off score for the 0–100 Oswestry Disability Index is approximately 10 points [42] or an 18% reduction of the pre-surgery score [62]; for the pain VAS approximately 20 points (on a 100-point scale) [42], and for the 0–24 point Roland Morris disability approximately 4 points [11, 62]. The minimal clinically relevant changes for generic health scales, such as the SF36, and other secondary outcome measures, such as psychological distress, have been less well investigated. However, these tend to be less responsive to surgery [10, 42] and often the minimal clinically relevant change borders on the value for the minimal detectable difference (i.e. 95% confidence intervals for the measurement error) for these instruments [42], rendering difficult the identification of ''real change'' as opposed to ''random error'' in a given individual.

Instruments used in the prediction of outcome

Psychometric testing with reliable and valid questionnaire instruments is a very efficient way to collect a great deal of information about the patient undergoing spinal surgery. Questionnaires provide a means of cross-checking clinical interview data. It is recommended that the latter be carried out prior to questionnaire assessment, as the trusting relationship built up during the clinical interview increases the chances of patients responding honestly to the questionnaires. Further, an adequate process of informed consent, in which patients are clearly informed why the questionnaire data are being collected and who will have access to them, increases the chances of patients giving truthful answers as opposed to answers they may perceive to be ''desirable.'' Nevertheless, it is a common finding that the information derived from the clinical interview and the psy-chometric testing is not always consistent, e.g. a patient not showing any concerns about surgery during the interview may actually reveal himself to be quite fearful as assessed by the psychometric testing. It is this additional information derived from the psychometric testing that renders it so useful. Other advantages include its objectivity, i.e. the results do not depend on the person who analyses the information, and its standardized nature. A knowledge of normative values for each questionnaire allows individual scores to be interpreted in relation to the scores of many others and hence provides clinicians with information as to whether an individual's scores are extreme, i.e. should be considered to be prognostic regarding the outcome or not. In acknowledgement of the evidence amassed in recent years from high quality research, decisions regarding treatment should consider the results of psychometric assessments.

Many of the instruments commonly used to assess risk factors for a poor surgical outcome have been discussed at length in other contributions in this special issue or in recent review articles. For instance, instruments assessing demographic and medical factors are discussed in the article by Häfeli and colleagues. The contribution by Elfering covers work-related instruments. Other recently published reviews have covered in detail the use of self-report questionnaires for current pain and pain history, disability, and general health status [47, 66]. Hence, the instruments that will be discussed in the present article are restricted to the most important individual psychosocial characteristics, i.e. personality characteristics, emotional antecedents and reaction to pain, and the way individuals attribute pain to certain work or physical activity factors.

This description of instruments is by no means exhaustive, but includes those that have consistently displayed prognostic value in relation to the outcome of spine surgery.

Personality is defined by the American Psychological Association as ''deeply ingrained patterns of behaviours, which include the way one relates to, perceives, and thinks about the environment and oneself.'' From this definition it is understandable that personality may influence individual reactions to pain and spinal surgery. The Minnesota Multiphasic Personality Inventory (MMPI) [43] and its revision, the MMPI-2, have been used in the area of chronic pain and surgery outcome for more than 50 years [9]. The core and most commonly used scales of the MMPI and MMPI-2 comprise three validity scales and ten clinical scales [32]. In the context of chronic pain and the outcome of spine surgery, two scales, hypochondriasis (Hs) and hysteria (Hy), have been shown to be valuable predictors of outcome [8]. Both scales were originally constructed to identify patients whose psychopathology is manifested in physical symptoms in

the absence of organic pathology. In spine surgery, however, where the indisputable presence of organic pathology is a pre-requisite for surgery, the psycho-analytical theory behind the scales is neglected and both scales are instead considered to address sensitivity to pain. In one study, it was shown that when Hs and Hy were assessed prior to discography of both disrupted and normal discs, patients with significant Hs and Hy scores ($T > 75$) were more likely to report pain regardless of the status of the discs injected [7, 8]. The MMPI takes 45–75 min to administer, some items are difficult to respond to, and the calculation of test scores is complicated. An alternative, the Maudsley Personality Inventory (MPI) [46], also known as the Eysenck Personality Questionnaire, comprises 80 items for three scales of introverted/extroverted personality, emotional stability (neurotic tendencies), and false discovery. The MPI takes around 10 min to complete and is available in many languages. The most recent and widely accepted approach to personality is the Big-Five factor model. The five-factor model of personality comprises the dimensions of neuroticism, extraversion, openness, agreeableness, and conscientiousness [19, 63]. Short measures include the 45-item bipolar adjective rating list developed by Ostendorf and colleagues [72, 73], which was further reduced to 30 items by Schallberger and Venetz [81]. The latter demonstrated that the reduced version is satisfactory in terms of factorial structure and internal consistencies of the scales. Each scale consists of six bipolar items on a six-point scale, with each pole ranging from "very" (1 and 6), "quite" (2 and 5), and "rather" (3 and 4). Principal components analysis with a forced five-factorial solution replicated the factorial structure, with 27 out of 30 items loading on the appropriate factor [26].

Patients who suffer from (chronic) pain and undergo surgery are often depressed. While depression, chronic pain, and surgical outcome seem to be linked, the direction of causal pathways is not clear. Depression appears as an antecedent and as a consequence of chronic pain in approximately equal measure (for further details, see later): Atkinson et al. showed depression to precede chronic pain in 42% of patients and to be a consequence in 58% [4]. The MMPI includes a depression subscale, but the instruments more commonly used to assess depression are the 21-item Beck Depression Inventory [5], the CES-D scale [78], and the 23-item Zung Depression Inventory [100]. All instruments display adequate psychometric properties and have been translated into several languages. Following the work of Greenough et al. [34], the modified Zung scale is now commonly combined with the 13-item Modified Somatic Perception Questionnaire (MSPQ) [58] by simple addition of the scores of each questionnaire to provide an accurate measure of psy-chological disturbance. The so-called Distress and Risk Assessment Method (DRAM) [60], developed in a sample of 567 chronic pain patients in Scotland, also includes the combination of both instruments. Completion of the 45-item instrument takes approximately 10 min. The DRAM is used to categorize patients into one of four groups: normal; at risk; distressed-depressive with high Zung scores but moderate MSPQ scores; and distressed-somatic with the reversed pattern of scores. It is considered to be a good screening tool for patients with risk of poor outcome in spinal surgery, although not all studies have been able to confirm its predictive power in all patient groups (see later). Depression has also been assessed using the Psychological General Well-Being Questionnaire [24]. This questionnaire consists of a total of 22 questions on the following six subscales: anxiety, depression, well-being, self-control, health, and vitality. Patients rate each question on a six-point Likert scale. The three-item measure of depression asks whether participants (a) felt depressed, (b) felt downhearted and blue, and (c) felt sad, discouraged, and hopeless during the past month.

Several studies have explored the validity and the predictive power of fear-avoidance beliefs (about activities of daily living and work) in relation to outcome. There is increasing evidence in the literature that the Fear-Avoidance Beliefs Questionnaire (FABQ) [96] is a valid and reliable tool that can be helpful in predicting treatment outcome in LBP and in surgically treated patients [85]. Pfingsten et al. [77] and Staerkle et al. [85] both reported the validation of a German version of the FABQ and Chaory et al. have validated it in French [17]. Further research is required to examine whether a splitting of the FABQ work scale into "work as a cause" and "work prognosis" factors, as proposed by Pfingsten et al. [77], is better in predicting surgical outcome than the single scale proposed by the original authors and supported in the German version of Staerkle et al. [85].

To some extent, the diversity of available instruments and the versions within instruments represents a considerable obstacle in attempting to compare results across studies. As such, some sort of standardization is highly recommended. This is the motivation behind current efforts to standardize both extended batteries of questionnaires (e.g. the Deutsche Schmerzfragebogen, DSF [25, 68, 90]), and the very brief "core" or screening sets for use in the clinical routine (e.g. [61, 95]).

Predictors of outcome of spinal surgery

The present review focuses specifically on predictors of the outcome of surgical treatment for spinal disorders, although there is some suggestion that there may be a

certain amount of overlap with the factors that determine the prognosis of non-operative treatment of LBP [18], especially in relation to the chronic stage and for some of the psychological attributes such as depression and fear-avoidance beliefs and certain medical factors such as the number of previous treatments undertaken.

In examining the literature on predictors of surgical outcome, we will focus mainly on studies carried out in the last 10–15 years (> 1990). Recent imaging modalities and operative techniques have advanced so much since the 1980s that negative explorations are now quite rare and the clinical presentation is more straightforward [16]; hence, studies using diagnostic techniques and/or operative methods that are no longer state-of-the-art may identify predictors that are of little relevance today.

A search of the literature reveals a plethora of studies in which predictor factors have been assessed. In many of these, the primary aim of the study was simply to report the outcomes for a given procedure, and the factors associated with a good or bad outcome were considered as incidental or supplementary information. The latter (often retrospective studies) tend to be less robust in terms of their scientific quality, as detailed in Sect. "Statistical methods used to identify predictors" earlier in this chapter. Other studies have specifically set out to examine prospectively the predictors of outcome for a given spinal disorder or surgical technique, and it is the results of these studies that are most helpful in identifying the variables that consistently emerge as predictors. Some of the recent key studies [prospective, multiple predictor variables (or new and interesting predictors) examined, multivariate analyses used, valid outcome instruments] are listed in Table 1; the list is, however, by no means exhaustive.

The most commonly examined predictors of surgical outcome can be loosely categorized into the following groups: biological/demographic, work-related, psychosocial, and medical (Table 1). In addition to these, and increasing in popularity as a relatively unexplored avenue for explaining some of the variance in outcomes, is the notion of the "patients' expectations of surgery" [57, 65].

As alluded to in the earlier sections, one must bear in mind a number of factors when examining the agreement between studies for the variables identified as "predictors." Firstly, predictors can only be found among the variables that are examined in the first place; and secondly, the failure to evaluate potentially important predictor variables in some studies can lead to overestimation of the importance of the variables that are examined or to emphasis being placed on different but closely related variables carrying similar information. Further, in studies of very small groups of patients, the sample sizes for different outcome groups may be too small (especially in relation to the size of the "poor outcome" group, which tends to contain just a minority

of patients) to sufficiently power the study and allow it to identify potentially relevant, real differences.

In most of the following studies only the statistical significance of the associations between outcome and risk factors was given, and, only rarely, the extent of the variance in outcome accounted for. The implications of this will be discussed again later.

Biological/demographic and health behaviour/lifestyle variables

Biological/demographic variables

Numerous retrospective studies have shown a negative association between the patient's age at surgery and outcome, although most of the prospective studies have shown no influence of age (see Table 1) or have even found improved outcomes in older patients (cervical spine) [75]. In part, the role of age may be explained by the outcome measure being investigated: where work issues are concerned, then it is more likely that older age at operation will result in less positive results with regard to return-to-work. It is also unclear in many studies (especially when bivariate analyses were used) whether the duration of symptoms was controlled for. The latter is one of the strongest predictors of a poor outcome (see later), and especially in chronic disorders tends to show a correlation with age. Hence, age may be acting in part as a marker for symptom duration, where the latter has not been simultaneously accounted for.

Gender is also highlighted by many retrospective studies as a potential predictor of outcome, although most prospective studies have failed to find such an association. Those that do, tend to show that men have a better outcome than women (see Table 1). An association with "maleness" is difficult to explain: postulated mechanisms include the notion of gender acting as an indirect marker for various (negative) psychological factors [93], biological differences in the healing potential of men and women, or (with respect to fusion) gender-related differences in the mechanical loading/muscle compressive forces promoting new bone growth [76].

Body weight has rarely been found to be a predictor of the outcome; many studies show no influence (Table 1), although one recent study showed obesity to have a negative effect on the outcome [8].

Health behavioural/lifestyle factors

Few studies have examined "health behavioural" or "lifestyle" factors as predictors of outcome, although it is conceivable that these could be important in determining an individual's response to major surgery. Intuitively, one might imagine that a higher level of pre-surgical

Table 1 Summary of recent prospective studies of predictors of outcome of spinal surgery, grouped according to outcome measure used (global score, back function pain, return to work). See text for details

Reference	L/C	Surgery, Indication	No. Pats	FU	B/M	Outcome	Result	More aged	Male gender	Smoking	High BMI/weight	Low income	Low education	Low job level	Worker's comp/disability	Heavy job	Long sick leave/unemployment	Job satis/stress/resignation	MMPI scales	Depression/psych distress	Family reinforcement	Pain drawings/somatic behaviour/symp	Copying strategies	Neuroticism	no. affected levels	Long duration symptoms	Severity, clinical	Severity, imaging	Comorbidity/self-rated low health	Previous ops	% Variance accounted for R2
Block 2001	L	laminect/ discect; fusion (cLBP)	204/259	>6mo	M	global score†	24% good, 42% fair	o			-				-	-			-		-		-								-
Carragee 2003	L	discectomy, herniation	180/187	>2y	M	global score# (& function, re-op rate)	mean 73% improvement	o	o						-	o	o									o	mix	+			-
Junge 1995	L	disc surgery (herniation/other)	328/381	1 y	M	global score (SC*)	52% good	o	o					-	-	o	-			o	-	o				-	mix	o			
Kohlboeck 2004	L	discectomy, herniation	48/58	6 mo	M	global score$	56% good	o	o																	-	+				
Nygaard 2000	L	microdiscectomy, herniation	132	1 y	M	global score°	-	o	o																	-					21%
Hagg 2003	L	fusion, degen cLBP	201/232	2y	M	global category	63% improved	o	o	o					o	o	o	o	o	o	o	o			o	o	o	mix			
Schade 1999	L	discectomy, herniation	42/46	2 y	M	global score (SC*)	74% exc/good				o							o			o	-			o	o	o	+			58%
Spratt 2004	L	decompression, stenosis	36/40	1 y	M	global score‡	58% successful															-			o						
Rothoerl 2002	L	discectomy, herniation	219	9.9 mo	B	global score (PROLO)	60% favourable			-																					
Carragee 1997	L	discectomy, herniation	48/51	>2 y	M	global score#	75% exc/good	o	+	o	o				-											o	o	+			75%
Woertgen 1997 ESJ	L	discectomy, herniation	113/121	1y	B	global (PROLO)	70% good			-	o															-	mix	+	-		
Hagg 2003 a, b	L	fusion, cLBP	201/232	2y	M	function (ODI)	25% improvement	o	o	o					o	o	o	o	-	o	o	-		-	o	o	o	mix	o		
McGregor 2002	L	decompression, stenosis	65/84	1 y	M	function (ODI)	20% improvement	o	o	o										o						o			o		11-50%
Ng 204	L	discectomy, herniated disc	103/113	1 y	M	function (ODI)	77% clin rel improvement			o							o									-		+			
Peolsson 2003	C	decompression & fusion, degen cNP	74/103	>12 mo	M	function (NDI)	30% NDI score = 20%	o	o	-									-	-						-	mix	o	-		23%
Schade 1999	L	discectomy, herniation	42/46	2 y	M	function (RM)	-	o			o		-		o		o	o		o	o				o	o	mix	+	o		46%
Solberg 2005	L	microdiscectomy, herniated disc	180/228	>1y	M	function (ODI) & pain	mean improvement 69%						-		o		-			o										o	27%
Trief 2000	L	67%fusion, 30% decomp; cLBP	102/150	6 & 12 mo	M	function (DPQ)	-								o	o															36-41%
Woertgen 1999	L	discectomy, herniation	98/121	3, 12, 28 mo	M	function (LBOS)	66% successful*	o	o	o																	mix				
Woertgen 1997a, b	L	discectomy, herniation	113/121	1y	B	function (LBOS, PROLO)	70-76% good	o	o	-																-	mix		-		

Table 1 (Contd.)

Study	L/C	Procedure	No. patients	FU	B/M	Outcome	Outcome description	FU%
Katz 1999	L	decompression, stenosis	199/272	2 y	M	pain, symptoms, satis, walk	37% mild/no pain; 73% satisfied	22-33%
McGregor 2002	L	decompression, stenosis	65/84	1 y	M	pain	mean decrease 30%	11 to 50%
Ng 2004	L	discectomy, herniated disc	103/113	1 y	M	pain	mean decrease 60%	-
Peolsson 2003	C	decompression & fusion, degen cNP	74/103	>1 y	M	pain	30% patients pain =10 (0-100 scale)	30%
Schade1999	L	discectomy, herniation	42/46	2 y	M	pain	83% complete relief leg pain	
Trief 2000	L	67%fusion, 30% decomp; cLBP	102/150	6 & 12 mo	M	pain	59-65% better	30%
Grebner 1999	L	nucleotomy (1st), herniation	82	6 mo	B	pain		-
Woertgen 1997	C	foraminotomy, herniation	50/54	1 y	B	pain	94% improved	
Hagg 2003	L	fusion, degen cLBP	201/232	2y	M	RTW	38% working	-
Kaptain 1999	C	discectomy, herniation	269/269	>10 mo	M	RTW	84% working	-
Schade 1999	L	discectomy, herniation	42/46	2 y	M	RTW	81% working fulltime	31%
Trief 2000	L	67%fusion, 30% decomp; cLBP	102/150	6 & 12 mo	M	RTW	51% working	-
Young 1997	L	microdiscect, DH	348	>10 mo	M	RTW	75% working	
Grebner 1999	L	nucleotomy (1st), herniation	82	6 mo	B	RTW		
Ng 2004	L	discectomy	103/113	1 y	M	satisfaction	65% exc/good	-
Peolsson 2004	C	decompression & fusion	74/103	>1 y	M	fusion	27% pseudoarthr	14%

+' = positive effect on outcome; '-' = negative effect on outcome; 0 = no effect on outcome; 'MIX' = some positive, some negative, some no effect

"L/C": L lumbar, C cervical; "No. patients" = number of patients followed-up out of original group; "B/M": B bivariate, M multivariate analyses used; "FU": follow-up duration
RTW, Return To Work; ODI Oswestry Disability Index; RM Roland Morris Disability scale; SC Stauffer-Coventry (pain, working, medication/physician visits); DPQ Dallas Pain Questionnaire
LBOS low back outcome score; PROLO Prolo Score; R2 = % variance accounted for by all listed predictors in the final multiple regression model
†pain, function, medication use; #pain, function, satisfaction, medication use; ‡pain, function, claudication; $Pain, function, RTW, qual of life; ° pain, clin exam, function, medication
* results differed slightly at different FU times, as did the predictors (only stable ones mentioned here)

physical fitness would allow a more rapid return to normal functioning after surgery. To the authors' knowledge, fitness or the participation in regular exercise has been examined in only one retrospective study [6] and was not found to be associated with outcome after percutaneous lumbar discectomy. Results from the authors' own studies suggest that the regular participation in exercise/physical activity for many years prior to the operation (but not necessarily exercise habits at the time of the intervention)—i.e. exercise as a "lifetime habit"—is significantly associated with a positive outcome after decompression surgery (unpublished observations).

Smoking is a relatively frequently examined predictor factor, especially in relation to the outcome after spinal fusion. In some studies it has been shown to have a negative impact on outcome, whereas in many others it has had no effect (Table 1). It has been suggested that tobacco use must be examined as a dose–response relationship in order to reveal associations that can be obscured by expressing it as a dichotomous variable (yes/no to a smoking habit) [54]. While the inhibitory effects of nicotine on fusion itself have been established [2, 31], it is also possible that smoking may simply reflect other factors—such as negative health behaviour (low physical activity levels, alcohol use), lower education/social level, manual job—and thereby act as a marker for these in determining the outcome. Interestingly, even in a subgroup of patients with no signs of pseudoarthrosis, smoking still predicted clinical outcome and return-to-work in patients undergoing fusion [31].

Work-related factors

Work-related predictors include such variables as worker's compensation, disability pension, work-status before surgery, duration of sick leave, and heaviness of the job.

The majority of studies that have examined the effect on outcome of the involvement in disability pension claims or worker's compensation issues have confirmed that these have a negative impact on the result of surgery, especially in relation to return-to-work or "global outcomes" (see Table 1 and also [21, 35, 36, 54, 55, 92]). In one large high quality study, however, workers' compensation showed no effect with the outcome in multivariate models [40]. The authors suggested that the strength of such an association may in part depend on the social insurance system in the given country [40]. One large retrospective study showed that while compensation status was predictive of the 2-year outcome after fusion, it no longer had any influence (in terms of back-specific function scores) after 10 years [74].

Although rarely examined in prospective studies, retrospective studies have shown that the involvement of a lawyer in compensation claims has a consistent negative predictive value for various outcomes after spinal fusion [21, 22, 54]. Cynics may interpret this finding as evidence for the premeditated instruction to magnify symptoms for the purposes of secondary gain; some studies have even shown that lawyers may advise their clients how to respond to psychological assessments in order to better their chances of success with their disability claims (see discussion in [54]). Others have suggested that litigious patients experience an increased somatic sensitivity to pain as a consequence of financial incentives and social–contextual variables [27].

Long pre-operative sick leave is a consistent negative predictor of return-to-work [40, 71, 89] and of global outcome, overall satisfaction or back-specific function [48, 82]. This highlights the importance of providing timely intervention once a clear-cut diagnosis that can be remedied by surgery has been made (see later).

Job heaviness (physically strenuous work) has been examined as an independent predictor in only a few studies, and the results appear to be somewhat conflicting: in one retrospective study on herniated disc patients, heavy manual work was a negative predictor of overall outcome and post-operative work-status 10 years after lumbar discectomy [56]. A prospective study of patients with chronic degenerative LBP revealed a similarly negative relationship in relation to outcome measured with a combined global score [8], whereas a further study on fusion patients [40] and two others on discectomy patients showed no influence of heavy work on the outcome [16, 98]. Intuitively, it may be expected that, while work-status may not necessarily govern the degree of pain and disability reported after surgery, it may well influence an individual's chances of returning to a job requiring the performance of heavy manual duties.

Psychological and sociological factors

Psychological factors

Psychological factors are one of the mostly commonly investigated predictors of surgical outcome, although their overall importance still remains equivocal and may be dependent on the spinal disorder in question [14].

Some of the early studies carried out in the 1980s showed slight to moderate associations between certain scales on the MMPI (most commonly Hs, Hy, depression, and admission of symptoms scales) and outcome after disc surgery/fusion. These studies encouraged the development of scoring systems, which included MMPI measures, to assist in predicting surgical outcome from various baseline indicators [8, 83, 91]. In view of the various psychometric and practical problems associated with use of the MMPI in pain patients [59], new or modified methods of assessing psychological

characteristics have been introduced, which focus primarily on the measurement of depression, anxiety, and/or heightened somatic awareness. More recently, other psychological characteristics have become of interest as potential predictor factors, such as coping strategies [8, 33], fear-avoidance beliefs (about work and physical activity) [85], and various workplace psychological factors (stress, satisfaction, "resigned" attitude, etc.) [80]. Overall, these have led to mixed results, in terms of their ability to reliably predict the outcome.

Using pain drawings and inappropriate signs, Greenough et al. [35, 36] reported in two retrospective studies that "psychological distress" was predictive of a poor outcome after anterior fusion. Van Susante et al. [93] used a "psychogenic back pain score" to examine prospectively the outcome after lumbosacral fusion of three types of patient group: organic, uncertain, and psychogenic. It was shown that the "organic" group had a much better outcome in terms of pain, disability, and medication use than did the "psychogenic" group. In patients undergoing discectomy, depression was found to be a significant predictor of global outcome [53, 80] and return-to-work [80]. A recent prospective study by Trief et al. [89] investigated the influence of baseline depression, state anxiety, somatic anxiety, and hostility on the outcome after lumbar spine surgery [mostly fusion (68%) and decompressive laminectomy (30%)]: using multivariate analyses, the DRAM, which classifies patients as either "normal," "at-risk," or "distressed," was found to be a significant predictor of outcome in terms of work-status, change in back pain and leg pain, and the "daily activities" and "work–leisure activities" scales of the Dallas Pain Index.

Junge et al. [48] found that certain aspects of pain behaviour (search for social support) were significantly associated with a poor global outcome in patients undergoing disc surgery; although depression did not show a significant association, there was a tendency for higher baseline values in patients with a poor outcome and depression was therefore included in the pre-screening tool developed by the group. In prospectively studying patients undergoing discectomy [44] or fusion [88], two studies failed to reproduce the findings of Trief et al. [89], in that the DRAM scores were found to have no predictive power in relation to back-function (Oswestry Disability Index). Similarly, neither depression [40] nor pain drawings [41] were able to predict outcome (any domain) after fusion for chronic LBP (Table 1). Greenough et al. [37] were also unable to reproduce their earlier findings [35] in a later retrospective study on patients undergoing posterolateral surgery. Notably, in all these studies, psychological disturbance was improved after surgery in patients with a good outcome. No association could be found between depression and outcome in studies on spinal stenosis patients undergoing decompression [51, 64].

In a large group of patients followed up after spinal surgery (for mixed diagnoses), Staerkle et al. [85] showed that fear-avoidance beliefs at baseline were a significant predictor of work-loss at 6 months. They uniquely explained 12% variance in the outcome, which was approximately one-third of the total variance explained by the whole predictive model (socio-demographic variables, pain variables, and fear-avoidance beliefs together explained 37% variance), representing a moderate effect size.

It has been suggested that the poor results of surgery reported in psychologically disturbed patients may reflect intervention in patients who did not have surgically remediable pathology [94], and this appears to have been verified by the maFny recent studies of Carragee (see [14]). This group has shown that patients with acute and subacute sciatica in association with a clearly identifiable, severe disc herniation have a very high chance of dramatic and lasting improvement with surgery and that standard psychometric tests in these patients fail to predict the outcome. Even severe emotional distress in patients who underwent early, appropriate surgical intervention did not correlate with adverse outcomes, although the same psychometric profile in patients with chronic sciatic pain and disability did predict worse outcomes compared with less emotionally distressed patients with the same level of chronicity. It was concluded that with prolonged pain and emotional distress, adverse and possibly self-perpetuating psychological and social changes may significantly decrease the impact of disc surgery [14].

All in all, and in view of the conflicting evidence, it would not appear prudent to recommend that patients with a surgically remediable pathology be denied surgery simply on the basis of their pre-operative psychological status. Nonetheless, it may be a useful strategy to identify patients with long-lasting symptoms and a high level of distress who might benefit from an additional psychological treatment before and/or accompanying surgical treatment; decreased levels of distress may then increase the impact of surgical treatment.

Sociological factors

Low social functioning (as measured with quality of life instruments) was identified as a significant negative predictor of re-operation rate in a retrospective study on fusion patients [30], and of global outcome, pain, and quality of life in a mixed group of spine-surgery patients [86].

In patients undergoing lumbar disc surgery, job level was found to be a significant predictor of combined global outcome [48]. An interesting study on military personnel undergoing cervical disc surgery showed that both position (rank) and duration of the individual's military career (but not economic forms of secondary gain per se) were significant predictors of return to active duty [50]. In

some studies, a low education level and/or low income have been shown to predict a negative surgical outcome in terms of either the total costs associated with workers' compensation [22], return-to-work [99], or global outcome/function [48, 56, 98]. It has been suggested that because individuals with a better education, a higher income, and at a higher level on the job-ladder tend to have greater responsibilities, personal investment may override the discomfort caused by any residual post-operative symptoms and encourage a return-to-work [50].

Occupational mental stress and job-related resignation have been shown to be negatively associated with return-to-work and post-operative pain relief/disability, respectively [80]. Job-related resignation reflects a "resigned" attitude to work-related troubles, job continuation despite dissatisfaction, the notion that the current situation must be accepted because things might otherwise be worse, and that expectations are limited as an employee [80]. The significance of the impact of job satisfaction on return-to-work is well documented in the back-pain literature [20, 95].

Social support from the spouse [80], search for social support (as a pain behaviour) [48], and family reinforcement of pain [8] have all been associated with a more negative outcome after surgery. It is suggested that this kind of "support"—in which relatives take over the patient's jobs or responsibilities, encourage rest, and provide more attention when the pain appears greatest [27]—serves to reinforce the illness status and thereby encourages the adoption of "passive" behaviour [27, 80].

Medical factors

Diagnosis-specific clinical factors

Few studies have been able to identify clinical variables that are predictive of outcome after spinal surgery. Hagg et al. [40] reported no significant predictive effect on the outcome after fusion of various baseline pain-provocation (flexion/extension), trunk flexibility, and neurological tests, with the exception of abnormal motor function, which was associated with a poorer outcome. One study has shown that pre-operative sensory deficit is associated with a good outcome (in terms of back-specific function), but the relationship was only evident 28 months after surgery and not at the 3 or 12 month follow-ups [98], suggesting it may have been a spurious finding. In the same study, the presence of a positive SLR test at $< 30°$ was associated with an unfavourable outcome at each time-point, and significantly so at 12 months. In contrast, Kohlboeck et al. [53] showed that, pre-operatively, the Lasegue sign was a good indicator of a successful outcome. Junge et al. considered the deficiency of reflexes to be predictive of a better outcome in their pre-screening instrument developed for disc surgery patients [48].

Imaging

The recent widespread use of the MRI scan in the assessment of spinal disorders has considerably improved the ability of surgeons to understand spinal pathology, especially in relation to disc herniation [14]. In two studies, Carragee and Kim showed that in patients with sciatica, the anterioposterior length of the herniated disc material and the ratio of disc area to canal area seen on MRI [15] as well as the degree of anular competence and type of herniation seen intraoperatively [16] had a stronger association with surgical outcome (pain, function, medication use, and satisfaction) than did any clinical or demographic variables. Other studies have shown that patients with an uncontained herniated disc had a better functional outcome 1 year after surgery than did those with a contained herniation [69]. Using multiple regression analysis of a range of medical variables (including MRI findings) and psychosocial variables, Schade et al. [80] reported that MRI-identified nerve-root compromise and the extent of herniation were the strongest independent predictors of global surgical outcome 2 years after surgery in patients undergoing lumbar discectomy. In contrast, return-to-work could not be predicted by any clinical or imaging variables and was instead determined by various psychosocial factors.

Sun et al. [87] retrospectively compared the outcome after adjacent two-level lumbar discectomy in patients with radicular pain attributable to nerve-root impingement either with or without concomitant osseous degenerative changes at the same level. The proportion of patients with an excellent/good global outcome (Mac-Nab classification) was significantly higher in the group with only a herniated disc (86%) compared with the group in which osseous changes were also present (57%).

One large study showed that low disc height (less than 50%) was one of the most significant positive predictors of outcome (back-specific function) in patients with degenerative chronic LBP undergoing spinal fusion [40]. In contrast, Peolsson et al. [75, 76] found that disc space narrowing was without any prognostic significance for functional outcome. In patients undergoing lumbar fusion, a surgical diagnostic severity score, based on pre-surgical imaging, had no predictive power for either disability status, global outcome, or physical or social functioning subscales of the SF20 [21].

In the study of Peolsson et al. [75, 76], pre-operative segmental kyphosis at the level to be operated was the strongest predictor of pain and disability 2 years after cervical decompression with fusion, although the proportion of the explained variance was low.

Pain history

A consistent predictor of poor outcome for various different diagnoses and types of outcome is the duration of

symptoms prior to the operation (Table 1). In studies that failed to identify this association, closely related variables (e.g. long-term sick leave, work-disability claim) were often chosen for inclusion in the multivariate model, especially in predicting return-to-work [40, 89].

Prior operations on the spine has been identified as a risk factor for poor outcome in a couple of studies [50, 64], although, interestingly, satisfaction with repeat operations is purportedly higher when there is a history of good results from previous operations and no epidural scarring requiring surgical lysis [70].

The number of affected (or operated) levels is often assumed to be negatively associated with the outcome, although only few (mostly retrospective) studies have actually demonstrated such a relationship with regards to disability status after fusion [21, 29, 50], the long-term clinical outcome after laminectomy [45], or the risk of requiring subsequent fusion after discectomy [87]. This relationship is believed by some to be related to resulting post-operative spinal instability [45]. A number of other studies on various diagnostic groups have been unable to confirm this association at all [1, 39, 76, 84]. Again, identifying the correct surgically treatable lesion(s) may be of greater importance; if this is not done, then increasingly poor results can obviously be expected as increasingly more levels are wrongly operated.

General medical

Many studies have shown that, especially in older populations of patients, poor general health in terms of other joint problems or systemic diseases (comorbidity) appears to have a significant negative influence on the outcome of spinal surgery [14, 48, 51]. However, some studies have failed to find any clear association [40, 84]. Perhaps the poor patient-rated outcomes in comorbid patients reflect, in part, a cross-contamination of the outcome instruments (especially those assessing function [67]), leading to an over-estimation of the true back-specific disability. Either way, it is important to make patients with comorbidity aware that the operation is being carried out for the specific spinal lesion identified and that it will not serve as a panacea for all their ongoing medical problems.

Surgery-related factors

All the factors assessed so far for their role in determining the outcome of surgery are somewhat "extrinsic" to the surgical procedure itself. The assumption tends to be that the surgeon himself is infallible and that the only reason for failure relates to inherent characteristics of the patient himself. Certainly, surgical skill is an aspect that is difficult to examine within the context of clinical trials, but we must concede that a certain proportion of failures are attributable not to the patient but to failure of the technique used, the hardware, and surgical complications. Furthermore, it is incumbent upon the surgeon to perform an accurate diagnostic work-up and to critically assess the indications for surgery; any shortcomings in this respect will naturally increase the potential for an unsatisfactory result. A recent study, in which the rates of surgery for herniated disc and spinal stenosis were compared across different spine service areas in the State of Maine (USA), found that the rates varied up to fourfold among the areas examined [52]. Interestingly, the outcomes for patients in the area with the lowest surgery-rate were significantly superior to those in the high surgery-rate areas (79% vs 60% with marked/complete pain relief, respectively) [52]. The patients in the higher-rate areas generally had less severe symptoms at baseline than did those in the lowest-rate area. The authors concluded that the variability may have been related to differences in physicians' preferences or thresholds for severity with regard to recommending an operation and their criteria for the selection of patients. Waddell et al. have argued that distress may increase the pressure for surgery, and that inappropriate symptoms and signs may obscure the physical assessment, leading to a mistaken diagnosis of a surgically treatable lesion [94]. In this instance, psychological factors may affect the outcome of surgery indirectly if inappropriate illness behaviour leads to inappropriate surgery [94].

As far as technical success is concerned, one of the most commonly assessed surgical outcomes is the achievement of arthrodesis after fusion surgery, although it has long been a matter of debate whether the presence of pseudoarthrosis has any influence on the subsequent patient-orientated outcome. Some studies have shown that pain relief in particular is greater when solid fusion is achieved [13, 76, 97], although it explains only a small proportion of the variance in pain outcome (4% [76]). In one recent study of interbody cage lumbar fusion, although 84% patients achieved solid fusion, only approximately 40–50% patients demonstrated a successful outcome in terms of pain, quality of life, global outcome, and work-disability status [54]. Other retrospective studies have indicated that the presence of radiological arthrodesis has no influence on either back-function [37, 74] or work-disability status [29] after fusion.

Summary

It is extremely difficult to identify unequivocal predictor factors that can be used to accurately predict the outcome of surgery. Many predictor factors are contentious, or are at least very specific to the patient profile, the diagnosis, and the surgical technique under investigation (and perhaps even to the institution in which the investigations were carried out). Moreover, the length

Table 2 Generally consistent predictors of poor-outcome (see also Table 1)

Long duration of symptoms
Severity of pathology on MRI (for disc herniation only)
Comorbidity/other joint problems/poor general health
Psychological distress (e.g. depression, anxiety),
  especially in patients with chronic pain
Family reinforcement of pain, especially in
  patients with chronic pain
Smoking (especially for fusion)
Job dissatisfaction/resignation
Worker's compensation
Long-term sick-leave/work disability

and type of follow-up appear to play such a decisive role, as does the scientific quality of the study in which the predictor factors are investigated, that it becomes almost impossible to provide a simple recipe for predicting the outcome of surgery with any certainty on an individual basis. Some predictor models or screening tools have been developed [8, 40, 48, 83], but few [49] have been investigated in a different patient group or under conditions that differ from those in which they were originally developed, limiting their applicability for general use. Moreover, the proportion of variance in outcome explained by even a combination of the strongest predictors is often relatively low, suggesting that we have a long way to go before being able to rest easily having refused someone surgery on the basis of unfavourable baseline characteristics.

In reality, the best that science can offer is a series of factors that can be considered to "influence" (rather than predict) the outcome of surgery, which might be put together in a list and considered together with the patient's diagnosis, the proposed operative technique, and the characteristics of the patient, in order to discuss with the patient reasons that might cause his outcome to deviate from the "optimal" or "normal" (Table 2). The opportunity (time), encouragement (education and positive messages), support and resources (referral to appropriate supporting services) to modify risk factors that are indeed modifiable can be offered, and realistic expectations can be discussed with the patient before the decision to operate is made. Such approaches have already proven worthwhile with respect to such factors as smoking cessation prior to fusion surgery [31]. Clear risk factors for a poor outcome in relation to return-to-work are long-term sick leave/receipt of disability benefit. Every effort should therefore be made to keep the individual in the workforce despite ongoing symptoms and plans for surgery. In patients with a particularly heavy job, consultation with occupational physicians to implement ergonomic change or provide job re-training to allow lighter duties might later ease the way back into the workplace. Patients with a degenerative condition and/or concomitant systemic or joint disease should be counselled that their condition is unlikely to return to normal and that only a small percentage of them will have complete pain relief or a complete return to pre-morbid function. Patients with long-lasting symptoms and a high level of distress may benefit from an additional psychological treatment, before and/or accompanying the surgical treatment.

These modifications, per se, might ultimately result in a greater satisfaction with surgery—if satisfaction is, indeed, determined by having had one's expectations fulfilled. Most spinal surgery is carried out for disorders that are not life-threatening, and while time may be of the essence for disorders with a very clear-cut diagnosis [69, 71, 79], there are also many that do not require immediate surgical treatment. This is not to suggest that a simple wait and see policy be adopted without further intervention; instead, active measures to minimize risk factors should be taken in order to best prepare the patient for a potential future surgical procedure, and evidence-based conservative treatments should be persevered with in the meantime. Recent studies suggest that many of the latter are as good as surgery for some of the less well-defined indications (e.g. chronic LBP due to degenerative changes) commonly dealt with by spinal fusion [12, 28] and these treatments may be worth considering as an alternative in patients for whom the outcome of surgery is uncertain.

References

1. Amundsen T, Weber H, Nordal HJ, Magnaes B, Abdelnoor M, Lilleas F (2000) Lumbar spinal stenosis: conservative or surgical management? A prospective 10-year study discussion 1435–1426. Spine 25:1424–1435
2. Andersen T, Christensen FB, Laursen M, Hoy K, Hansen ES, Bunger C (2001) Smoking as a predictor of negative outcome in lumbar spinal fusion. Spine 26:2623–2628
3. Asch HL, Lewis PJ, Moreland DB, Egnatchik JG, Yu YJ, Clabeaux DE, Hyland AH (2002) Prospective multiple outcomes study of outpatient lumbar microdiscectomy: should 75 to 80% success rates be the norm?. J Neurosurg Spine 96:34–44
4. Atkinson JH, Slater MA, Patterson TL, Grant I, Garfin SR (1991) Prevalence, onset, and risk of psychiatric disorders in men with chronic low back pain: a controlled study. Pain 45:111–121
5. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J (1961) An inventory for measuring depression. Arch Gen Psychiatry 4:561–571
6. Bernd L, Schiltenwolf M, Mau H, Schindele S (1997) No indications for percutaneous lumbar discectomy?. Int Orthop 21:164–168
7. Block AR, Vanharanta H, Ohnmeiss DD, Guyer RD (1996) Discographic pain report Influence of psychological factors. Spine 21:334–338

8. Block AR, Ohnmeiss DD, Guyer RD, Rashbaum RF, Hochschuler SH (2001) The use of presurgical psychological screening to predict the outcome of spine surgery. Spine J 1:274–282

9. Block AR, Gatchel RJ, Deardorff WW, Guyer RD (2003) The psychology of spine surgery. American Psychological Association, Washington

10. Bombardier C (2000) Outcome assessments in the evaluation of treatment of spinal disorders. Spine 25:3100–3103

11. Bombardier C, Hayden J, Beaton DE (2001) Minimal clinically important difference Low back pain: outcome measures. Pain 28:431–438

12. Brox JI, Sorensen R, Friis A, Nygaard O, Indahl A, Keller A, Ingebrigtsen T, Eriksen HR, Holm I, Koller AK, Riise R, Reikeras O (2003) Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. Spine 28:1913–1921

13. Buttermann GR, Garvey TA, Hunt AF, Transfeldt EE, Bradford DS, Boachie-Adjei O, Ogilvie JW (1998) Lumbar fusion results related to diagnosis. Spine 23:116–127

14. Carragee EJ (2001) Psychological screening in the surgical treatment of lumbar disc herniation. Clin J Pain 17:215–219

15. Carragee EJ, Kim DH (1997) A prospective analysis of magnetic resonance imaging findings in patients with sciatica and lumbar disc herniation Correlation of outcomes with disc fragment and canal morphology. Spine 22:1650–1660

16. Carragee EJ, Han MY, Suen PW, Kim D (2003) Clinical outcomes after lumbar discectomy for sciatica: the effects of fragment type and anular competence. J Bone Joint Surg Am 85-A:102–108

17. Chaory K, Fayad F, Rannou F, Lefevre-Colau MM, Fermanian J, Revel M, Poiraudeau S (2004) Validation of the French version of the fear avoidance belief questionnaire. Spine 29:908–913

18. COST B13 Action (2004) Guidelines for the management of chronic low back pain. www.backpaineurope.org

19. Costa PTJ, McCrae RR (1985) The NEO Personality Inventory manual. Psychological Assessment Resources, Odessa

20. Coste J, Delecoeuillerie G, Cohen de Lara A, Le Parc JM, Paolaggi JB (1994) Clinical course and prognostic factors in acute low back pain: an inception cohort study in primary care practice. BMJ 308:577–580

21. DeBerard MS, Masters KS, Colledge AL, Schleusener RL, Schlegel JD (2001) Outcomes of posterolateral lumbar fusion in Utah patients receiving workers' compensation: a retrospective cohort study discussion 747. Spine 26:738–746

22. DeBerard MS, Masters KS, Colledge AL, Holmes EB (2003) Presurgical biopsychosocial variables predict medical and compensation costs of lumbar fusion in Utah workers' compensation patients. Spine J 3:420–429

23. Deyo R, Centor RM (1986) Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. J Chronic Dis 39 (11):897–906

24. Dupuy HJ (1984) The Psychological General Well-Being (PGWB) Index. In: Assessment of quality of life in clinical trials of cardiovascular therapies. Le Jacq, New York, pp 170–183

25. Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N, Carr DB, Chandler J, Cowan P, Dionne R, Galer BS, Hertz S, Jadad AR, Kramer LD, Manning DC, Martin S, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robbins W, Robinson JP, Rothman M, Royal MA, Simon L, Stauffer JW, Stein W, Tollett J, Wernicke J, Witter J (2005) Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 113:9–19

26. Elfering A, Kälin W, Semmer NK (2000) Stability and change in job satisfaction at the transition from vocational training into "real work". Swiss J Psychol 59:256–271

27. Epker J, Block AR (2001) Presurgical psychological screening in back pain patients: a review. Clin J Pain 17:200–205

28. Fairbank JHF, Frost H, Wilson-MacDonald J, Yu LM, Barker K, Collins R (2005) Spine Stabilisation Trial Group: Randomised controlled trial to compare surgical stabilisation of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain. The MRC Spine Stabilisation trial BMJ 330(7502):1233

29. Franklin GM, Haug J, Heyer NJ, McKeefrey SP, Picciano JF (1994) Outcome of lumbar fusion in Washington State workers' compensation. Spine 19:1897–1903

30. Glassman SD, Dimar JR, Johnson JR, Minkow R (1998) Preoperative SF-36 responses as a predictor of reoperation following lumbar fusion. Orthopedics 21:1201–1203

31. Glassman SD, Anagnost SC, Parker A, Burke D, Johnson JR, Dimar JR (2000) The effect of cigarette smoking and smoking cessation on spinal fusion. Spine 25:2608–2615

32. Graham JR (1990) The MMPI-2: assessing personality and psychopathology. Oxford University Press, New York

33. Grebner M, Breme K, Rothoerl R, Woertgen C, Hartmann A, Thome C (1999) Coping and convalescence course after lumbar disk operations. Schmerz 13:19–30

34. Greenough CG, Fraser RD (1991) Comparison of eight psychometric instruments in unselected patients with back pain. Spine 16:1068–1074

35. Greenough CG, Taylor LJ, Fraser RD (1994a) Anterior lumbar fusion A comparison of noncompensation patients with compensation patients. Clin Orthop 30–37

36. Greenough CG, Taylor LJ, Fraser RD (1994b) Anterior lumbar fusion: results, assessment techniques and prognostic factors. Eur Spine J 3:225–230

37. Greenough CG, Peterson MD, Hadlow S, Fraser RD (1998) Instrumented posterolateral lumbar fusion Results and comparison with anterior interbody fusion. Spine 23:479–486

38. Grob D, Benini A, Junge A, Mannion AF (2005) Clinical experience with the Dynesys semirigid fixation system for the lumbar spine: surgical and patient-orientated outcome in 50 cases after an average of 2 years. Spine 30:324–331

39. Gunzburg R, Keller TS, Szpalski M, Vandeputte K, Spratt KF (2003) Clinical and psychofunctional measures of conservative decompression surgery for lumbar spinal stenosis: a prospective cohort study. Eur Spine J 12:197–204

40. Hagg O, Fritzell P, Ekselius L, Nordwall A (2003a) 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

41. Hagg O, Fritzell P, Hedlund R, Moller H, Ekselius L, Nordwall A (2003b) 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

42. Hagg O, Fritzell P, Nordwall A, Group SLSS (2003c) The clinical importance of changes in outcome scores after treatment for chronic low back pain. Eur Spine J 12:12–20

43. Hathaway SR, McKinley JC (1951) The Minnesota Personality Inventory manual revised. The Psychological Corporation, New York

44. Hobby JL, Lutchman LN, Powell JM, Sharp DJ (2001) The distress and risk assessment method (DRAM). J Bone Joint Surg Br 83:19–21

45. Iguchi T, Kurihara A, Nakayama J, Sato K, Kurosaka M, Yamasaki K (2000) Minimum 10-year outcome of decompressive laminectomy for degenerative lumbar spinal stenosis. Spine 25:1754–1759

46. Jensen AR (1958) The Maudsley Personality Inventory. Acta Psychol 14:314–325

47. Junge A, Mannion AF (2004) Questionnaires for patients with back pain Diagnosis and outcome assessment. Orthopade 33:545–552

48. Junge A, Dvorak J, Ahrens S (1995) Predictors of bad and good outcomes of lumbar disc surgery A prospective clinical study with recommendations for screening to avoid bad outcomes. Spine 20:460–468

49. Junge A, Frohlich M, Ahrens S, Hasenbring M, Sandler A, Grob D, Dvorak J (1996) Predictors of bad and good outcome of lumbar spine surgery A prospective clinical study with 2 years' follow up discussion 1064–1055. Spine 21:1056–1064

50. Kaptain GJ, Shaffrey CI, Alden TD, Young JN, Laws ER Jr, Whitehill R (1999) Secondary gain influences the outcome of lumbar but not cervical disc surgery discussion 223–215. Surg Neurol 52:217–223

51. Katz JN, Stucki G, Lipson SJ, Fossel AH, Grobler LJ, Weinstein JN (1999) Predictors of surgical outcome in degenerative lumbar spinal stenosis. Spine 24:2229–2233

52. Keller RB, Atlas SJ, Soule DN, Singer DE, Deyo RA (1999) Relationship between rates and outcomes of operative treatment for lumbar disc herniation and spinal stenosis. J Bone Joint Surg Am 81:752–762

53. Kohlboeck G, Greimel KV, Piotrowski WP, Leibetseder M, Krombholz-Reindl M, Neuhofer R, Schmid A, Klinger R (2004) Prognosis of multifactorial outcome in lumbar discectomy: a prospective longitudinal study investigating patients with disc prolapse. Clin J Pain 20:455–461

54. Lacaille RA, Deberard MS, Masters KS, Colledge AL, Bacon W (2005) Presurgical biopsychosocial factors predict multidimensional patient: outcomes of interbody cage lumbar fusion. Spine J 5:71–78

55. Little DG, MacDonald D (1994) The use of the percentage change in Oswestry Disability Index score as an outcome measure in lumbar spinal surgery. Spine 19:2139–2143

56. Loupasis GA, Stamos K, Katonis PG, Sapkas G, Korres DS, Hartofilakidis G (1999) Seven- to 20-year outcome of lumbar discectomy. Spine 24:2313–2317

57. Lutz GK, Butzlaff ME, Atlas SJ, Keller RB, Singer DE, Deyo RA (1999) The relation between expectations and outcomes in surgery for sciatica. J Gen Intern Med 14:740–744

58. Main CJ (1983) The Modified Somatic Perception Questionnaire (MSPQ). J Psychosom Res 27:503–514

59. Main CJ, Spanswick CC (1995) Personality assessment and the Minnesota Multiphasic Personality Inventory 50 years on: do we still need our security blanket?. Pain Forum 4:90–96

60. Main CJ, Wood PLR, Hollis S, Spanswick CC, Waddell G (1992) The distress and risk assessment method—a simple patient classification to identify distress and evaluate the risk of poor outcome. Spine 17:42–52

61. Mannion AF, Elfering A, Staerkle R, Junge A, Grob D, Semmer NK, Jacobshagen N, Dvorak J, Boos N (2005) Outcome assessment in low back pain: how low can you go? Eur Spine J (in press)

62. Mannion AF, Junge A, Grob D, Dvorak J, Fairbank JCT (2005) Development of a German version of the Oswestry Low Back Index. Part 2: Sensitivity to change after spinal surgery. Eur Spine J (in press)

63. McCrae RR, John OP (1992) An introduction to the five-factor model and its applications. J Pers 60:175–215

64. McGregor AH, Hughes SP (2002) The evaluation of the surgical management of nerve root compression in patients with low back pain Part 1: The assessment of outcome. Spine 27:1465–1470

65. McGregor AH, Hughes SP (2002) The evaluation of the surgical management of nerve root compression in patients with low back pain Part 2: Patient expectations and satisfaction discussion 1476–1477. Spine 27:1471–1476

66. Muller U, Duetz MS, Roeder C, Greenough CG (2004) Condition-specific outcome measures for low back pain Part I: Validation. Eur Spine J 13:301–313

67. Muller U, Roeder C, Dubs L, Duetz MS, Greenough CG (2004) Condition-specific outcome measures for low back pain Part II: Scale construction. Eur Spine J 13:314–324

68. Nagel B, Gerbershagen HU, Lindena G, Pfingsten M (2002) Development and evaluation of the multidimensional German pain questionnaire. Schmerz 16:263–270

69. Ng LC, Sell P (2004) Predictive value of the duration of sciatica for lumbar discectomy A prospective cohort study. J Bone Joint Surg Br 86:546–549

70. North RB, Campbell JN, James CS, Conover-Walker MK, Wang H, Piantadosi S, Rybock JD, Long DM (1991) Failed back surgery syndrome: 5-year follow-up in 102 patients undergoing repeated operation discussion 681–690. Neurosurgery 28:685–690

71. Nygaard OP, Kloster R, Solberg T (2000) Duration of leg pain as a predictor of outcome after surgery for lumbar disc herniation: a prospective cohort study with 1-year follow up. J Neurosurg Spine 92:131–134

72. Ostendorf F (1990) Sprache und Persönlichkeitsstruktur: Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit [Language and personality structure: the validity of the five-factor model of personality]. Roderer, Regensburg

73. Ostendorf F, Angleitner A (1992) On the generality and comprehensiveness of the five-factor model of personality. Evidence for five robust factors in questionnaire data. In: Modern personality psychology. Critical reviews and new directions. Harvester Wheatsheaf, New York, pp 73–109

74. Penta M, Fraser RD (1997) Anterior lumbar interbody fusion A minimum 10-year follow-up. Spine 22:2429–2434

75. Peolsson A, Hedlund R, Vavruch L, Oberg B (2003) Predictive factors for the outcome of anterior cervical decompression and fusion. Eur Spine J 12:274–280

76. Peolsson A, Hedlund R, Vavruch L (2004) Prediction of fusion and importance of radiological variables for the outcome of anterior cervical decompression and fusion. Eur Spine J 13:229–234

77. Pfingsten M, Kroner-Herwig B, Leibing E, Kronshage U (2000) Validation of the German version of the Fear-Avoidance Beliefs Questionnaire (FABQ). Eur J Pain 4:259–266

78. Radloff LS (1977) The CES-D Scale: a self-report depression-scale for research in the general population. Appl Psychol Meas 1:385–401

79. Rothoerl RD, Woertgen C, Brawanski A (2002) When should conservative treatment for lumbar disc herniation be ceased and surgery considered? Neurosurg Rev 25:162–165

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

81. Schallberger U, Venetz M (1999) Kurzversionen des MRS-Inventars von Ostendorf (1990) zur Erfassung der fünf "grossen" Persönlichkeitsfaktoren [Brief versions of Ostendorf's MRS inventory for the assessment of the Big-Five personality factors]. Universität Zürich, Zürich: Berichte aus der Abteilung Angewandte Psychologie 30:1–51

82. Solberg TK, Nygaard OP, Sjaavik K, Hofoss D, Ingebrigtsen T (2005) The risk of "getting worse" after lumbar microdiscectomy. Eur Spine J 14:49–54

83. Spengler DM, Ouellette EA, Battie M, Zeh J (1990) Elective discectomy for herniation of a lumbar disc Additional experience with an objective method. J Bone Joint Surg Am 72:230–237

84. Spratt KF, Keller TS, Szpalski M, Vandeputte K, Gunzburg R (2004) A predictive model for outcome after conservative decompression surgery for lumbar spinal stenosis. Eur Spine J 13:14–21

85. Staerkle R, Mannion AF, Elfering A, Junge A, Semmer NK, Jacobshagen N, Grob D, Dvorak J, Boos N (2004) Longitudinal validation of the fear-avoidance beliefs questionnaire (FABQ) in a Swiss-German sample of low back pain patients. Eur Spine J 13:332–340

86. Stärkle R, Mannion AF, Junge A, Elfering A, Grob D, Dvorak J, Boos N (2002) The influence of baseline psychological factors on outcome after spine surgery. SIROT, San Diego

87. Sun EC, Wang JC, Endow K, Delamarter RB (2004) Adjacent two-level lumbar discectomy: outcome and SF-36 functional assessment. Spine 29:E22–E27

88. Tandon V, Campbell F, Ross ER (1999) Posterior lumbar interbody fusion Association between disability and psychological disturbance in noncompensation patients. Spine 24:1833–1838

89. Trief PM, Grant W, Fredrickson B (2000) A prospective study of psychological predictors of lumbar surgery outcome. Spine 25:2616–2621

90. Turk DC, Dworkin RH, Allen RR, Bellamy N, Brandenburg N, Carr DB, Cleeland C, Dionne R, Farrar JT, Galer BS, Hewitt DJ, Jadad AR, Katz NP, Kramer LD, Manning DC, McCormick CG, McDermott MP, McGrath P, Quessy S, Rappaport BA, Robinson JP, Royal MA, Simon L, Stauffer JW, Stein W, Tollett JJW (2003) Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. Pain 106:337–345

91. Uomoto JM, Turner JA, Herron LD (1988) Use of the MMPI and MCMI in predicting outcome of lumbar laminectomy. J Clin Psychol 44:191–197

92. Vaccaro AR, Ring D, Scuderi G, Cohen DS, Garfin SR (1997) Predictors of outcome in patients with chronic back pain and low-grade spondylolisthesis discussion 2035. Spine 22:2030–2034

93. Van Susante J, Van de Schaaf D, Pavlov P (1998) Psychological distress deteriorates the subjective outcome of lumbosacral fusion A prospective study. Acta Orthop Belg 64:371–377

94. Waddell G, Morris EW, Di Paola MP, Bircher M, Finlayson D (1986) A concept of illness tested as an improved basis for surgical decisions in low-back disorders. Spine 11:712–719

95. Waddell G, Burton AK, Main CJ (2003) Screening to identify people at risk of long-term incapacity for work A conceptual and scientific review. Royal Society of Medicine Press, London

96. Waddell G, Newton M, Henderson I, Somerville D, Main CJ (1993) A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain 52:157–168

97. Wetzel FT, McCracken L, Robbins RA, Lahey DM, Carnegie M, Phillips FM (2001) Temporal stability of the Minnesota Multiphasic Personality Inventory (MMPI) in patients undergoing lumbar fusion: a poor predictor of surgical outcome. Am J Orthop 30:469–474

98. Woertgen C, Rothoerl RD, Breme K, Altmeppen J, Holzschuh M, Brawanski A (1999) Variability of outcome after lumbar disc surgery. Spine 24:807–811

99. Young JN, Shaffrey CI, Laws ER Jr, Lovell LR (1997) Lumbar disc surgery in a fixed compensation population: a model for influence of secondary gain on surgical outcome. Surg Neurol 48:552–558

100. Zung WW (1965) A Self-Rating Depression Scale. Arch Gen Psychiatry 12:63–70