Björn Zoëga
Johan Kärrholm
Bengt Lind

Outcome scores in degenerative cervical disc surgery

Received: 26 April 1999
Revised: 13 September 1999
Accepted: 22 December 1999

This study was supported with grants from: the Gothenburg Medical Society, the Greta and Einars Askers Foundation, and Gothenburg University.

B. Zoëga (✉) · J. Kärrholm · B. Lind
Department of Orthopaedics,
Sahlgrenska University Hospital,
41345 Gothenburg, Sweden
e-mail: bjorn.zoega@orthop.gu.se,
Tel.: +46-31-3421000,
Fax: +46-31-823584

Abstract Forty-six consecutive patients with neck pain and arm radiculopathy were treated with anterior cervical discectomy and fusion. All patients had neurological symptoms corresponding to a herniated disc and/or spondylosis at one or two cervical levels, verified by magnetic resonance imaging. The patients were stabilized with an anterior graft and randomized to either fixation with a CSLP plate or no internal fixation. Preoperatively and 2 years postoperatively the patients filled in a questionnaire that included a modified Million Index, a modified Oswestry Index and the Zung Depression Scale. They were also asked to register their pain in the arm and in the neck on a vertical visual analogue scale (VAS). At the 2-year follow-up, an unbiased observer graded the patients' clinical outcome using Odom's criteria. A test-retest procedure was carried out to examine the questionnaire reproducibility. In the group that was operated at one level, there was no significant improvement in any of the scores. Nevertheless, 81% of the patients were satisfied with the outcome of the surgery. All scores improved in the group operated at two levels. The pain in the

neck and arm, as measured on a VAS, decreased in both groups. The improvement in arm pain was significantly more pronounced in patients operated with a plate at two levels compared to those who were operated without a plate. At the 2-year follow-up, patients with an excellent or good result according to Odom's criteria had a lower Million Index ($P < 0.0005$), Oswestry Index ($P < 0.0005$), and Zung ($P = 0.024$) score, than the group classified as fair or poor. There was a significant correlation ($P < 0.0001$ for all scores) between the test and retest results. We conclude that the modified Million Index and Oswestry Index are clinically useful tools in the evaluation of outcome after degenerative cervical disc surgery. The clinical benefits of plate fixation were minimal. The outcome after surgery, measured with the Oswestry Index, Million Index and VAS for arm and neck pain, seems to correlate well with the classification of outcome by Odom.

Key words Anterior cervical discectomy and fusion · Cervical disc herniation · Plate fixation · Outcome scores

Introduction

Surgical treatment of cervical spondylosis and herniated discs with anterior discectomy and interbody fusion has

been reported to relieve pain and to improve function [1, 3, 14, 19]. The criteria according to Odom or Robinson are often used to measure the outcome of these procedures [6, 13, 16]. These predominantly subjective rating scales

may vary depending on the investigator, which obstructs comparisons between different studies and treatment alternatives. The lack of objective instruments also jeopardizes accurate determination of the speed of recovery and the true response to a treatment. The more objective measurements available, e.g. the visual analogue pain scale (VAS), only rate the pain and not the global function.

In contrast, there are several tests to assess the outcome of treatment of lumbar spine disorders [7, 9, 18]. Two commonly used measures are the Million Visual Analogue Disability Scale (Million Index [12]) and The Oswestry Low Back Pain Disability Questionnaire (Oswestry Index [8]). A modified version of the Oswestry Index, called the Neck Disability Index (NDI), has been used to assess the outcome of whiplash injuries [17], but so far this index has not become widely accepted.

The results after treatment of both cervical and lumbar spinal disorders are influenced by the patient's psychosocial background. This has been addressed in patients with lumbar spine disorders by psychometric tools such as the Zung Depression Scale [22]. A corresponding evaluation of the influence of psychosocial factors in disorders of the cervical spine is rarely, if ever done.

To prove that surgical treatment of degenerative diseases of the cervical spine is beneficial for the patient, the use of objective and validated instruments is mandatory. In this study we applied a modified Million Index, the Oswestry Index and the Zung Depression Scale to evaluate the outcome of cervical spine degenerative disc surgery.

Materials and methods

Between January 1994 and October 1995, 46 consecutive patients were referred to our department with neck pain and arm radiculopathy, and were subsequently scheduled for surgical treatment. All patients had neurological symptoms (not only radicular pain) corresponding to a herniated disc and/or spondylosis at one or two cervical levels, verified by magnetic resonance imaging. Forty-three patients were on sick leave preoperatively and three were receiving a medical pension due to an unrelated disease. All but two of the patients were taking analgesics on a daily basis. In this study, failure of conservative treatment or increasing neurological symptoms were considered to be an indication for surgical treatment. The patients were operated with anterior disc excision and

grafting using autologous bone from the left iliac crest (Smith Robinson technique [15]).

All patients included in the study were stabilized with an anterior graft and randomized to either fixation with a CSLP plate ($n = 24$, Synthes, Switzerland) or no internal fixation at all ($n = 22$). Randomization was done using sealed envelopes the day before surgery. Twenty-seven patients were operated at one level (1 at C4-C5, 12 at C5-C6 and 14 at C6-C7), 15 with plate fixation, and 12 without plate fixation. Nineteen patients were operated at two levels (5 at C4-C6 and 14 at C5-C7), nine with and ten without plate fixation (Table 1).

Surgical technique

All the patients were operated by the same two surgeons. The cervical spine was approached through an anterior transverse incision on the left side of the neck. The anterior longitudinal ligament was excised over the disc space, and the anterior half of the disc was removed macroscopically. The remaining posterior part of the nucleus pulposus and the herniated fragment was removed under magnification with a microscope. The graft was harvested from the left anterior iliac crest. Postoperatively, patients operated without a plate were treated with a Philadelphia collar for 6 weeks and those operated with plate fixation wore a soft collar for 6 weeks.

Clinical evaluation

Preoperatively, the patients filled in a questionnaire that included the Million Index, the Oswestry Index and the Zung Depression Scale. They were also asked to register their pain in the arm and in the neck on vertical visual analogue scales (VAS). The patients were asked to record their present pain, and their maximum and minimum pain during the last 3 days. The mean of these three scores was recorded. Two years postoperatively, the patients received the same questionnaires by mail, 3-5 days before their follow-up visit to the hospital. At this visit, the patients were asked to fill in a new copy of the same questionnaire (test-retest) and to register any remaining pain on the VAS scale. The patients were asked whether they were satisfied with the results of the surgery or not.

A neurologist without knowledge about the results of the randomization examined the patients both pre- and postoperatively. At the 2-year follow-up, he graded the patients' clinical outcome using Odom's criteria [13] (Table 2).

The Million Visual Analogue Disability Scale (Million Index)

This test was introduced in 1982 [12] and consists of 15 questions describing pain and disability, with responses expressed on a hori-

Table 1 Demographic data of the patients included in the study

	Plate		No plate	
	One level	Two level	One level	Two level
No. of patients	15	9	12	10
Age: median (range)	42 (25-60)	43 (25-51)	40 (27-54)	45 (27-57)
Sex (male/female)	7/8	4/5	8/4	3/7
Smokers	8	2	1	3
Duration of symptoms, months: median (range)	14 (1-80)	36 (1-180)	16 (8-36)	38 (6-120)

Table 2 Odom's criteria [13]

Outcome	Criteria
Excellent	All preoperative symptoms relieved. Abnormal findings improved
Good	Minimal persistence of preoperative symptoms. Abnormal findings unchanged or improved
Fair	Definite relief of some preoperative symptoms. Other symptoms unchanged or slightly improved
Poor	Symptoms and signs unchanged or worse

zontal analogue line. In the present study, small modifications of the wording were done to make the questions applicable to the neck region. The scores can vary between 0 (no disability) and 10 (invalid). The Million Index is the mean score for all answered questions.

The Oswestry Low Back Pain Disability Questionnaire (Oswestry Disability Index)

This test, designed in 1980 [8], is a multiple-choice questionnaire to be answered by the patient. It consists of ten sections relating to different activities of daily living. Each section describes six levels of disability associated with daily living, corresponding to score 0 (no disability) to 5 (invalid). The Oswestry Disability Index is calculated by dividing the total score by the number of sections answered. Multiplication by 20 gives a score from 0 to 100% disability. This score was designed to evaluate low back pain patients, but we are of the opinion that the same activities also are applicable to patients with neck and arm pain.

The Zung Depression Scale

This test was introduced in 1965 to evaluate patients with "clinical depression" [22]. It is a self-rating depression scale consisting of 20 items; ten cover positive and ten negative symptoms. When using the scale, the patients are asked to rate each of the 20 items as to how it applied to them at the time of testing, in four quantitative terms: a little of the time, some of the time, a large part of the time, and most of the time. In scoring the test, a value of between 1 and 4 is assigned to a response, depending on whether the item was worded positively or negatively. The total score can vary between 20 and 80. According to Zung, scores of 55 or higher indicate depression [22].

Statistics

Changes in the test results for the four groups between the pre- and the 2-year postoperative examination were evaluated using a Wilcoxon signed rank test. Because of the multiple comparisons made, the *P*-values were doubled (Bonferroni correction). To assess the results of the intervention, the concept of "effect size", was used, which was calculated by the formula: (mean preop. score–mean postop. score)/SD of preop. score [4, 11]. The effect size calculates the extent of change measured in a standardized way, which allows comparison between questionnaires. An effect size of 1.0 is equivalent to a change of one standard deviation in the sample. An effect size of less than 0.2 is small, 0.5 is moderate and above 0.8 is large [4, 11]. Differences between those operated with and those without plate fixation were calculated using the Mann-Whitney U test, using the type of Bonferroni correction described above.

The reproducibility (test-retest reliability) at 2 years was calculated using Spearman's correlation coefficient, which is a non-parametric correlation coefficient. The coefficient of reliability was computed by the standard error of the difference in scores between test 1 and test 2, multiplied by 1.96. Ninety-five percent of differences are expected to lie within two standard deviations [2]. To evaluate the validity of Odom's criteria, these were condensed into two groups (excellent or good and fair or poor), which were compared regarding the values according to the three different scoring systems and the recorded values on the VAS (construct validity) Logistic regression was used to analyse whether any of the preoperative variables (Tables 1,3) could be used as a predictor of outcome according to Odom's criteria. "Excellent" and "good" were classified as a satisfying outcome and "fair" and "poor" an unsatisfying outcome.

Results

No patients were lost to follow-up. One patient (one level, with plate fixation) developed pseudarthrosis and the plate broke between the 3rd and the 6th month. This patient was reoperated at 1 year, and the fusion healed. This patient's outcome was included in the results. No other implant-related complications were noted.

Fusion at one level

In this group, there was no significant improvement in any of the scores. Nevertheless, 81% of the patients (10 patients without plate fixation, 12 with plate fixation) were satisfied with the outcome of the surgery. Neither was there any improvement in the pain in the neck, whereas the pain in the arm tended to decrease in those operated without a plate ($P = 0.046$, Table 4) [20].

Fusion at two levels

A general tendency to improvement in all scores was noted, but the power of the analysis was only sufficient to demonstrate *P*-values less than 0.05 with regard to the Oswestry Index and the Million Index, and only in cases operated without a plate (Table 3). The pain in the neck and arm measured on a VAS decreased in both groups ($P = 0.016$ – 0.034 , Table 4). The improvement in arm pain was more pronounced in patients operated with a plate compared to those operated without a plate ($P = 0.02$) [21]. All the patients were satisfied with the outcome of their surgery.

There were no significant differences in Odom's criteria between those operated with and those without plate fixation (one level, $P = 0.8$; two level $P = 0.5$).

Construct validity

At the 2-year follow-up, patients with an excellent or good result according to the Odom's criteria had lower

Table 3 Median (range) values of Million Index, Oswestry Index and Zung scores, and distribution of patients by Odom's criteria for clinical outcome

	Plate		No plate	
	One level	Two levels	One level	Two levels
Million index				
Preop. score	5.6 0.6–8.3	7.0 4.2–8.7	5.0 2.0–7.1	6.5 4.2–7.7
2-years' postop. score	4.9 (1.5–8.2)	3.8 (0.3–8.5)	4.8 (0.2–7.6)	4.2 (0.3–7.0)
<i>P</i> -value ^a	> 0.3	0.056	> 0.3	0.026
Effect ^b				1.97
Oswestry Index				
Preop. score	42 (14–68)	52 (30–64)	36 (6–64)	48 (24–74)
2-years' postop. score	48 (8–76)	30 (0–66)	38 (0–64)	33 (0–62)
<i>P</i> -value ^a	> 0.3	0.075	> 0.3	0.016
Effect ^b				0.83
Zung score				
Preop. score	41 (27–61)	45 (31–55)	39 (30–66)	45 (26–65)
2-years' postop. score	46 (31–63)	35 (26–50)	49 (22–65)	37 (22–59)
<i>P</i> -value ^a	> 0.3	0.054	> 0.3	0.062
Effect ^b				
Odom's criteria				
Excellent/good	11	8	8	7
Fair/poor	4	1	4	3

^aThe *P*-values for the difference between the preoperative and 2-year follow-up are given. Wilcoxon rank sum test with Bonferroni correction was used

^bThe effect size [4, 11] is only calculated if the differences are significant

Table 4 Median (range) values of VAS for arm pain and for neck pain

	Plate		No plate	
	One level	Two levels	One level	Two levels
VAS arm pain				
Preop. score	4.0 (1.0–9.5)	5.1 (3.1–8.6)	6.3 (2.9–8.6)	5.8 (3.7–7.8)
2-years' postop. score	4.5 (0.3–8.0)	0.5 (0–3.7)	5.9 (0–7.8)	3.0 (0.6–5.7)
<i>P</i> -value ^a	> 0.3	0.016	0.046	0.034
Effect ^b		2.19	0.82	1.66
VAS neck pain				
Preop. score	5.4 (3.1–8.8)	6.4 (3.7–8.3)	6.3 (4.4–9.1)	6.3 (3.3–9.9)
2-years' postop. score	5.8 (1.7–8.8)	2.7 (0–6.6)	5.6 (0–8.2)	3.6 (0.5–7.1)
<i>P</i> -value ^a	> 0.3	0.024	> 0.3	0.034
Effect ^b		1.27		1.27

^aThe *P*-values for the difference between the preoperative and 2-years follow-up are given. Wilcoxon rank sum test with Bonferroni correction was used

^bThe effect size [4, 11] is only calculated if the differences are significant

Million ($P < 0.0005$), Oswestry ($P < 0.0005$), and Zung ($P = 0.024$) scores, than the group classified as fair or poor (Table 5). This former group also had less pain in the neck ($P < 0.0005$) and the arm ($P = 0.002$).

Reproducibility

There was a high and significant ($P < 0.0001$ for all scores) correlation between the test and retest results (Table 6). The mean difference (between test 1 and test 2) was 0.12 (SD 0.73; $P = 0.14$, Wilcoxon signed ranks test) for the Million Index, and for the Oswestry Index 0.33 (SD 5.1; $P = 0.84$). Although a high and significant correlation was found between the two Zung tests (mean difference 1.5,

SD 3.2; 95% CI of the mean = 0.31–2.3), there was a significant difference between the two tests ($P = 0.006$, Wilcoxon signed rank test). The differences between test 1 and test 2 proved to be normally distributed (Shapiro-Wilk's statistic = 0.95, $P = 0.13$). The coefficient of reliability for the Million Index was 0.2, for the Oswestry Index 1.5 and for the Zung score 1.0.

Prediction of outcome

Logistic regression analysis based on patient factors (gender, age, duration of pain, whiplash injury, preoperative Million Index, Oswestry Index and Zung score, and preoperative VAS scores for arm and neck pain), type of op-

Table 5 Distribution of different scores and arm and neck pain (VAS), presented as median (range), among patients classified according to Odom's criteria as excellent/good and fair/poor

	Excellent/good	Fair/poor	P-value
Million Index			
Preop. score	5.5 (0.6–8.7)	6.4 (4.9–8.3)	0.324
2-years' postop. score	3.6 (0.2–7.6)	7.2 (5.5–8.5)	< 0.0005
Oswestry Index			
Preop. score	37 (6–66)	48 (34–74)	0.154
2-years' postop. score	30 (0–32)	60 (34–76)	< 0.0005
Zung score			
Preop. score	41 (27–65)	45 (26–66)	0.528
2-years' postop. score	38 (22–63)	52 (32–65)	0.024
VAS, arm pain			
Preop. score	4.9 (1.0–9.5)	6.7 (3.2–8.2)	0.068
2-years' postop. score	2.7 (0–7.8)	5.9 (0.1–8.0)	0.002
VAS, neck pain			
Preop. score	5.5 (3.1–9.9)	6.9 (4.5–8.3)	0.120
2-years' postop. score	3.6 (0–8.0)	6.8 (4.2–8.8)	< 0.0005

^aMann-Whitney U test with Bonferroni correction

Table 6 The results (median and range) of the test-retest procedure at the 2-year follow-up

	Median (range)	r-value ^a
Million Index		
Test 1	4.4 (0.1–8.3)	0.93
Test 2	4.6 (0.2–8.5)	
Oswestry Index		
Test 1	36 (0–78)	0.95
Test 2	34 (0–76)	
Zung score		
Test 1	44 (21–64)	0.96
Test 2	42 (22–65)	

^aSpearman's ρ correlation coefficient

eration (plate fixation, number of levels fused) revealed that none of these variables had a significant influence on the outcome.

Discussion

Surgery is a generally accepted method to treat degenerated cervical discs with radiculopathy. Nevertheless, there are no generally accepted and validated methods that can be used to analyse the outcome of surgical against conservative treatment modalities. Such measures are also needed to obtain comparative data for assessing outcome of vari-

ous forms of surgery. Reproducible and unbiased outcome measurements have, however, rarely been used.

The method of measurement may influence the apparent results of treatment. Howe and Frymoyer [10] compared 14 ways of measuring outcome after treatment of low back pain. They found that the proportion of "successes" varied between 60 and 97%, depending on the outcome measure used. Subjective evaluations, e.g. the patient's own opinion of the operation, gave a higher proportion of "successful" results when compared to objective measures, such as return to original employment.

All outcome measure tools should be tested for their validity, reliability and responsiveness. To our knowledge, no outcome measures applicable to cervical spine disorders have been tested with regard to these criteria. The concept of effect size is one way to address the question of outcome. Some effects of a treatment may be statistically significant, but the magnitude of the improvement may be so small that it is clinically irrelevant. A treatment should be considered effective only if the improvement is statistically significant and the effect size is large. For those scores in the present study that had improved significantly between the preoperative and the 2-year follow-up investigation, the effect size was large.

We noted in this study that simple measurements like the VAS seem to register similar improvements to the composite scores such as the Million Index and the Oswestry Index. It is known that composite scores are liable to combine the variations of all the individual variables included in them, thus making the scores uncertain. For a score to be optimal, all included items should independently describe symptoms and functions related to the disease. A serious difficulty with VAS is the temporal variability of pain in cervical spine disorders. Many patients have periods of time, which may last from a few hours to many months, when they are almost without pain, while suffering attacks of severe pain in between these periods. This problem can be addressed either by asking the patient to comment only on present pain, or by asking the patient to recall the intensity of pain during attacks and duration of attacks. We chose to combine these methods, even though the latter is potentially unreliable due to the patient's potentially poor recall.

The Oswestry Index was devised for the measurement of disability and impairment in a low back pain population, and this index has previously been validated. The index is based on scores that are self-reported and is thus affected by the patients own perception of pain and desire to undertake the activities indicated in the index. It does not determine outcome by documentation of employment or more complex social or sporting activities. In the present study, the reduction in disability recorded with the Oswestry Index seems to correlate well with the outcome measured using Odom's criteria. The same is also true for the Million Index. Odom's criteria are widely accepted as the standard for clinical evaluation after cervical degener-

ative disc surgery, and are extensively used to report outcome after this type of surgery. Thus, the modified Oswestry and Million indexes used in the study also seem to be useful for measuring disability in cervical spine patients. The Zung Depression Scale has been shown to be a reliable tool for measuring psychological disturbance in patients with lumbar spine pain. It has a good specificity and sensitivity [9].

Patients operated at two levels in the present study had a significant improvement in many of the outcome scores tested. On the other hand, no or minor improvements were recorded in patients operated at one level, according to Odom's criteria, by which 70% were classified as excellent or good. The reason for this could be that the power of the tests was too small, due to a small number of cases. Another explanation might be that patients operated at one level had significantly lower preoperative Million and Oswestry scores than those operated at two levels. One can therefore speculate that those operated at two levels had a greater disability preoperatively than those operated at one level, as suggested by the preoperative scores, and that it may be more difficult to get a significant improvement for those with a lower preoperative disability. Despite this, most of the patients were satisfied with the outcome of their surgery. There were more scoring variables that improved significantly in the patients operated at two levels without plate fixation than those with plate fixation, but this difference is uncertain, due to the small number of patients in the two groups. The validity of the outcome measurement tools in the present study was not tested, but all of them correlated to each other and to the most commonly used (but also unvalidated) outcome measure (Odom's criteria), suggesting that these scores include relevant information. In this study, the Zung scores were not abnormal, indicating absence of depression. Nevertheless, they tended to decrease in patients operated at two levels, and remained almost unchanged in the one-level group. It

was found that patients with the worst outcome clinically measured with Odom's criteria also had a higher Zung score 2 years postoperatively. One explanation for this could be that people with certain psychological traits may be more likely to mention minor episodes of pain than other people. Alternatively, people with these psychological traits may have an increased sensitivity to cervical discomfort, and therefore experience more pain due to minor stimuli. Finally, there is a recognized association between chronic pain and depression.

A reliable test should give a consistent score on different occasions if the specific function that is measured has not changed. The test-retest procedure in this study resulted in a high correlation for all three tests, which confirms the stability of the tests. This was also confirmed with a relatively low coefficient of reliability. The Zung score showed a small, but significant, change within a short period of time. This could reflect the fact that the questions linked to statements in the Zung Depression Scale are ambiguous. It could also reflect a change in the patient's opinion resulting either from their previous confrontation with the Scale or from a time change in mental status. Because the difference between the test-retest procedures was small and the correlation between the tests was high, we think that the Zung score can still be used with confidence.

In conclusion, the modified Million Index and Oswestry Index are clinically useful tools in the evaluation of outcome after degenerative cervical disc surgery. All the tests administered at 2 years showed significant correlation with each other, and we can therefore conclude that they most likely measure the same condition. The outcome after surgery measured with Oswestry Index, Million Index and VAS for arm and neck pain, seems to correlate well with the more subjective classification of outcome according to Odom.

References

1. Aronson N, Filtzer DL, Bagan M (1968) Anterior cervical fusion by the Smith-Robinson approach. *J Neurosurg* 29: 396–404
2. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 307: 310
3. Bohlman HA, Emery SE, Goodfellow DB, Jones PK (1993) Robinson anterior cervical discectomy and arthrodesis for cervical radiculopathy. *J Bone Joint Surg Am* 75: 1298–1307
4. Cohen J (1977) Statistical power analysis for the behavioral sciences. Academic Press, New York
5. Dawson J, Fitzpatrick R, Murray D, Carr A (1998) Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br* 80: 63–69
6. DePalma AF, Rothman RH, Lewin-neck RE, et al (1972) Anterior interbody fusion for severe cervical disc degeneration. *Surg Gynecol Obstet* 134: 755–758
7. Deyo RA (1988) Measuring the functional status of patients with low back pain. *Arch Phys Med Rehabil* 69: 1044–1053
8. Fairbank JCT, Couper J, Davies JB, O'Brien JP (1980) The Oswestry low back pain disability questionnaire. *Physiotherapy* 66: 271–273
9. Greenough CG (1993) Recovery from low back pain. *Acta Orthop Scand* 48 [Suppl 254]: 1–34
10. Howe J, Frymoyer JW (1985) The effects of questionnaire design on the determination of end results in lumbar surgery. *Spine* 10: 804–805
11. Kazis LE, Anderson JJ, Meenan RF (1989) Effect sizes for interpreting changes in health status. *Med Care* [Suppl 3] 27: 178–179

-
12. Million R, Hall W, Haavik Nilsen K, Baker RD, Jayson MIV (1982) Assessment of the progress of the back-pain patient. *Spine* 7:204–212
 13. Odom GL, Finney W, Woodhall B (1958) Cervical disc lesions. *JAMA* 166:23–28
 14. Riley L, Robinson R, Johnson K, et al (1969) The results of anterior interbody fusion of the cervical spine. *J Neurosurg* 30:127
 15. Robinson RA, Smith GW (1955) Anterolateral cervical disc removal and interbody fusion for the cervical disc syndrome. *Bull Johns Hopkins Hosp* 96:223–224
 16. Robinson R, Walker A, Ferlic D (1962) The results of anterior interbody fusion of the cervical spine. *J Bone Joint Surg Am* 44:1569–1587
 17. Vernon H, Mior S (1991) The Neck Disability Index: a study of reliability and validity. *J Manipulative Physiol Ther* 14:409–415
 18. Waddel G, Main CJ (1984) Assessment of severity in low back disorders. *Spine* 9:204–208
 19. White A, Southwick W, Deponte R, et al (1973) Relief of pain by anterior cervical spine fusion for spondylosis. *J Bone Joint Surg Am* 55:525–534
 20. Zoëga B, Kärrholm J, Lind B (1998) One-level cervical spine fusion. A randomized study, with or without plate fixation, using radiostereometry in 27 patients. *Acta Orthop Scand* 69:363–368
 21. Zoëga B, Kärrholm J, Lind B (1998) Plate fixation adds stability to 2 level anterior fusion in the cervical spine: a randomized study using radiostereometry. *Eur Spine J* 7:302–307
 22. Zung WWK (1965) A Self-rating Depression Scale. *Arch Gen Psychiatry* 12:63–70