2-1-2004

Accurate Identification of adverse outcomes after cervical spine surgery

Charles C. Edwards II
Washington University Medical School in St. Louis

Yekaterina Karpitskaya
Washington University Medical School in St. Louis

Chuck Cha
Emory University School of Medicine

John G. Heller
Emory University School of Medicine

Carl Lauryssen
Washington University Medical School in St. Louis

See next page for additional authors

Authors
Charles C. Edwards II, Yekaterina Karpitskaya, Chuck Cha, John G. Heller, Carl Lauryssen, S. Timothy Yoon, and K. Daniel Riew

Accurate Identification of Adverse Outcomes After Cervical Spine Surgery

By Charles C. Edwards II, MD, Yekaterina Karpitskaya, MD, Chuck Cha, MD, John G. Heller, MD, Carl Lauryssen, MD, S. Timothy Yoon, MD, and K. Daniel Riew, MD

Investigation performed at the Department of Orthopaedic Surgery, Barnes-Jewish Hospital at Washington University School of Medicine, St. Louis, Missouri

Background: Retrospective clinical studies frequently utilize surgeon records as a source of outcomes data. The accuracy of data derived from surgeon records, however, is unknown. The purpose of the present study was to evaluate the accuracy of surgeon records in documenting the prevalence of subjective adverse outcomes.

Methods: Consecutive patients who had undergone anterior cervical arthrodesis by four spine surgeons during a ten-month period were included. Surgeon records from the routine six-week, three-month, and six-month postoperative visits were examined for documentation of persistent dysphagia and dysphonia. Patients completed surveys inquiring about the presence and magnitude of symptoms at these three time-points. Agreement between the surgeon records and the patient surveys was analyzed with use of the kappa coefficient.

Results: One hundred and sixty-six patients had 342 postoperative visits. Dysphagia was documented twenty-six times in the surgeon records, compared with 107 times on the patient surveys. Dysphagia was thus underreported in 80% of cases. Similarly, dysphonia was documented ten times in the surgeon records, compared with seventy-two times on the patient surveys. Poor correlation between the surgeon records and the patient surveys was observed regardless of symptom severity, previous anterior cervical surgery, anterior arthrodesis of three motion segments or more, arthrodesis cephalad to the fifth cervical level, and anterior cervical plate use. Poor correlation between the surgeon records and the patient surveys also was observed for each surgeon, regardless of subspecialty or institution.

Conclusions: Correlation between the surgeon records and the patient surveys was consistently poor, regardless of the specific patient and surgeon factor analyzed. While we chose to study dysphonia and dysphagia, it is conceivable that the results may be generalizable to many situations in which office notes are utilized to ascertain the prevalence of subjective adverse outcomes. These results suggest that the prevalence of such outcomes may be seriously underreported in studies that rely on the retrospective analysis of surgeon records.

Accurate knowledge of the occurrence of adverse outcomes after operative procedures is of critical importance to surgeons and patients. The majority of peer-reviewed clinical studies continue to be retrospective in design. A review of all spine-related clinical reports with data on complications that were published during 2000 in *Spine* and the American volume of *The Journal of Bone and Joint Surgery* revealed that forty-three (74%) of fifty-eight studies were retrospective in design. Of the forty-three retrospective studies, thirty-seven (86%) relied on existing medical records for the identification of specific postoperative complications. Only three studies (7%) used independent methods such as a symptom-focused survey (two studies) or clinical evaluation by an alternate healthcare provider (one study) to identify specific adverse outcomes.

Although surgeon records are widely utilized for the identification of adverse outcomes, they are subject to multiple potential sources of error, including underreporting by patients, underappreciation by surgeons, and lack of documentation by surgeons. Unfortunately, the effect of these potential biases and the accuracy of surgeon records as a source of data regarding subjective adverse outcomes are unknown.

The purpose of the present study was to investigate the accuracy of surgeon office notes as a source of data regarding subjective adverse outcomes after spine surgery. Specifically, we set out to test two hypotheses. First, we theorized that the occurrence of specific adverse outcomes after spine surgery as recorded in surgeon notes correlates poorly with the occurrence of such outcomes as reported by patients. Second, we theorized that poor correlation between surgeon notes and patient experience is a consistent phenomenon regardless of the symptom being evaluated and regardless of factors related

to the patient, surgeon, or surgical technique.

To test these hypotheses, patients who had undergone a commonly performed spinal procedure (anterior cervical arthrodesis) were evaluated. Two well-described adverse outcomes of this procedure (dysphagia and dysphonia) were arbitrarily selected as test symptoms. The occurrence of these symptoms according to blinded surgeon records and symptom-focused patient surveys was evaluated at six weeks, three months, and six months after surgery.

Materials and Methods

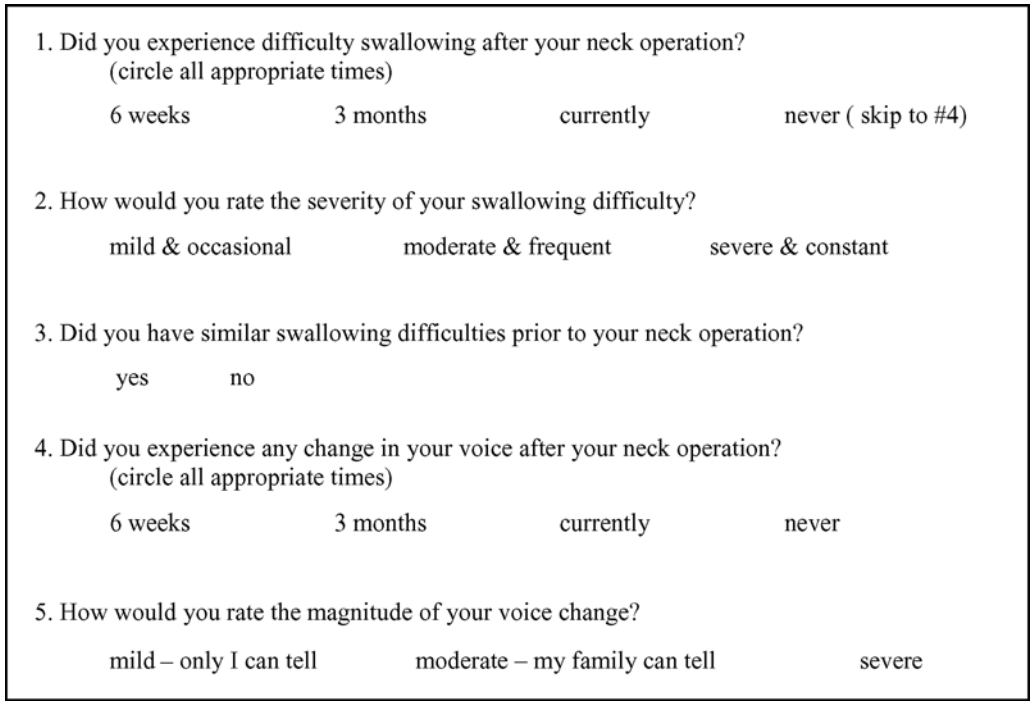
Four fellowship-trained spine surgeons (K.D.R., C.L., J.G.H., and S.T.Y.) who had no knowledge of the details of this study were recruited in October 2001. Approval for the study was obtained from the appropriate Human Investigations Committee. The medical records of 187 consecutive patients who had undergone anterior cervical spine arthrodesis and who had returned for at least one follow-up visit between January 1, 2001, and October 15, 2001, were examined. Any reference to the presence of dysphagia or dysphonia in the surgeons' notes at each follow-up time-period was recorded. If no reference to dysphagia or dysphonia was made in a particular note, then the symptoms were regarded as "absent."

A symptom-focused survey was mailed to the 187 eligible patients (Fig. 1). In this survey, patients were asked five questions related to the presence of dysphagia or dysphonia before and after the operation. Patients were asked about the severity of symptoms, if present, and whether such symptoms were present at six weeks, three months, and six months after surgery. When surveys were not returned by mail, the survey was conducted by means of a telephone interview by an independent clinician (Y.K.).

Survey data were obtained from 168 of the 187 eligible patients. Of the nineteen patients without survey data, fifteen patients could not be located and four refused to participate. Two patients who reported the occurrence of swallowing problems prior to surgery were also excluded.

The remaining 166 patients with surgeon notes and survey data comprised the sample population for this study. The study group included eighty-six male patients and eighty female patients with a mean age of 51.5 years (range, sixteen to eighty-two years). One hundred and twenty-three patients had undergone an anterior procedure only, and forty-three had undergone anterior and posterior procedures. A left-sided anterior cervical approach had been used for 125 patients (75%), and a right-sided approach had been used for forty-one (25%). Thirty-nine patients (23%) had undergone previous anterior cervical arthrodesis procedures. Anterior cervical plates had been used in 155 patients (93%). The number of motion segments that had been arthrodesed during the procedure was one for sixty-five patients (39%), two for fifty-eight (35%), three for thirty-four (20%), four for seven (4%), and five for two (1%). In seventy-four patients (45%), the arthrodesis had been performed cephalad to the fifth cervical level. The 166 patients returned for a total of 342 postoperative visits (163 six-week visits, 118 three-month visits, and sixty-one six-month visits) between January 1, 2001, and October 15, 2001.

For the purpose of data analysis, each clinic visit was regarded as a separate event. Patients had a minimum of one clinic visit (at six weeks) and a maximum of three visits (at six


1. Did you experience difficulty swallowing after your neck operation?
(circle all appropriate times)

6 weeks 3 months currently never (skip to #4)

2. How would you rate the severity of your swallowing difficulty?

mild & occasional moderate & frequent severe & constant

3. Did you have similar swallowing difficulties prior to your neck operation?

yes no

4. Did you experience any change in your voice after your neck operation?
(circle all appropriate times)

6 weeks 3 months currently never

5. How would you rate the magnitude of your voice change?

mild – only I can tell moderate – my family can tell severe

Fig. 1

Subjective patient survey regarding dysphagia and dysphonia.

TABLE I Prevalence of Dysphagia and Dysphonia as Recorded in Surgeon Office Notes and a Symptom-Focused Patient Survey*

	Symptom Recorded (Survey/Chart)				
	Yes/Yes	Yes/No	No/Yes	No/No	Kappa Coefficient
Dysphonia	6	66	4	266	0.09
Dysphagia†	14	93	12	222	0.10
Total	20	159	16	488	0.10

*Based on 342 patient visits. †A response regarding dysphasia was not provided by one patient on the follow-up questionnaire.

weeks, three months, and six months) within the study period. For each follow-up visit, two comparisons—one for dysphagia and one for dysphonia—were made between the surgeon notes and the patient surveys. The total number of comparisons between the surgeon notes and the patient surveys, therefore, varied for each patient (from two to six) depending on the number of follow-up visits that the patient had had within the study interval. Concordance between the surgeon notes and the patient surveys was evaluated for each symptom at each time-period. Statistical agreement was evaluated with use of the kappa coefficient, with a value of <0.25 corresponding with poor agreement.

Results

The prevalence of dysphagia and dysphonia varied depending on the data source. Over the three time-points considered, dysphagia was reported twenty-six times in the surgeon records and 107 times on the patient surveys (Table I). Dysphonia was reported ten times in the surgeon records and seventy-two times on the patient surveys. The patient surveys and surgeon records were concordant in documenting the presence of symptoms in only twenty instances. The presence of symptoms was reported on the patient surveys alone in 160 instances and in the surgeon records alone in sixteen instances. The agreement between the surgeon records and the patient surveys in documenting the presence of symptoms was poor for both dysphagia (kappa = 0.10) and dysphonia (kappa = 0.09).

For the purpose of analysis, the "true prevalence" of symptoms was defined as the presence of symptoms as documented in either the surgeon records or the patient surveys (Table II). On the basis of the surgeon records, nineteen patients (11%) experienced dysphagia and eight (5%) experienced dysphonia at one or more of the three postoperative time-points. On the basis of the patient surveys, ninety-five patients (57%) experienced dysphagia and forty-nine (30%) experienced dysphonia at one or more of the three postoperative time-points. On the basis of the number of patients with symptoms as documented in either the surgeon records or the patient surveys, the true prevalence of dysphagia was 57% (ninety-five of 166) and the true prevalence of dysphonia was 30% (fifty of 166). On the basis of the true prevalence, surgeon records underreported dysphagia by 80% (76/95) and dysphonia by 84% (42/50).

The level of agreement between the two data sources and the degree of underreporting on the surgeon records were evaluated at each of the three time-points. Poor agreement between the surgeon records and the patient surveys and a high degree of underreporting on the surgeon records was encountered at six weeks (kappa = 0.06; degree of underreporting = 83%), three months (kappa = 0.20; degree of underreporting = 76%), and six months (kappa = 0.09; degree of underreporting = 87%).

The severity of symptoms was reported on the patient surveys as mild, moderate, or severe. Analysis according to symptom severity revealed poor agreement and a high degree of underreporting for mild symptoms (kappa = 0.08; degree of underreporting = 87%) and moderate symptoms (kappa = 0.06; degree of underreporting = 86%). The results for severe symptoms were somewhat better but remained fair to poor (kappa = 0.18; degree of underreporting = 57%) (Table III).

The effect of potential risk factors for dysphagia and dysphonia on surgeon documentation was evaluated. Poor agreement between the surgeon records and the patient surveys was observed in association with each of the risk factors evaluated, including prior anterior cervical surgery (kappa = 0.13, degree of underreporting = 80%), anterior arthrodesis of three motion segments or more (kappa = 0.12, degree of underreporting = 77%), arthrodesis cephalad to the fifth cervical level (kappa = 0.14, degree of underreporting = 78%), circumferential procedures (kappa = 0.10, degree of underre-

TABLE II "True Prevalence" of Dysphagia and Dysphonia*

	Symptom Recorded†			
	Surgeon Records	Patient Survey	Surgeon Records or Patient Survey ("True Prevalence")	Underreporting on Surgeon Records
Dysphonia	8 (5%)	49 (30%)	50 (30%)	84% (42/50)
Dysphagia	19 (11%)	95 (57%)	95 (57%)	80% (76/95)

*Based on 166 patients. †The data are given as the number of patients, with the percentage in parentheses.

TABLE III Agreement of a Symptom-Focused Patient Survey and Surgeon Office Notes Based on Severity of Symptoms*

Severity†	Sample Size	Symptom Recorded (Survey/Chart)			Underreporting of Symptom on Chart	Kappa Coefficient
		Yes/Yes	Yes/No	No/Yes		
Mild	70	7	61	2	87%	0.08
Moderate	96	8	83	5	86%	0.06
Severe	23	5	13	5	57%	0.18

*Based on 342 patient visits. †Symptom severity was defined by the patients for both dysphagia and dysphonia.

porting = 81%), and anterior cervical plate use (kappa = 0.09, degree of underreporting = 83%).

The number of patients contributed to the study by the four surgeons varied, with surgeon A contributing 101 patients, surgeon B contributing thirty-four, surgeon C contributing twenty-two, surgeon D contributing nine. The level of agreement and the degree of underreporting were disappointing for surgeon A (kappa = 0.07, degree of underreporting = 87%), surgeon B (kappa = 0.23, degree of underreporting = 70%), surgeon C (kappa = 0.11, degree of underreporting = 90%), and surgeon D (kappa = 0.11, degree of underreporting = 57%). When the results were examined as a function of the surgeons' institutions, the level of agreement (kappa = 0.10 compared with 0.11) and the degree of underreporting on surgeon records (83% compared with 84%) were nearly identical.

Discussion

Clinical studies in which surgeon office notes are used to define the prevalence of adverse outcomes may be susceptible to numerous sources of error. The accuracy of surgeon notes as a source of data on adverse outcomes after spine surgery has not been well defined. The purpose of the current study was to define the accuracy of surgeon notes as a source of data on subjective adverse outcomes after spine surgery.

In this series of 342 office visits, agreement between the surgeon notes and the patient surveys was consistently poor, regardless of symptom severity, postoperative visit, surgeon identity, or institution. Even among patients considered to be at increased risk for postoperative dysphagia or dysphonia, the surgeon notes underreported the presence of symptoms in a large majority of cases.

A review of the literature on anterior cervical spine procedures revealed a broad range in the reported prevalence of postoperative dysphagia and dysphonia (from 1% to 60%) (Table IV). The reported prevalence of postoperative dysphonia and dysphagia seems to have a bimodal distribution, with some studies reporting a prevalence from 1% to 15% and other studies reporting a prevalence from 40% to 60%. A review of these studies indicated that those with a relatively low prevalence of symptoms involved the use of surgeon records as a source of data on adverse outcomes whereas those with a relatively high prevalence of symptoms involved the use of patient surveys.

TABLE IV Prevalence of Dysphagia and Dysphonia After Anterior Cervical Spine Procedures in Published Studies

Authors	Number of Patients	Prevalence of Dysphagia	Prevalence of Dysphonia	Data Source	
				Records	Survey
Winslow et al.[7] (2001)	497	60%	51%		X
Current study	166	57%	30%		X
Stewart et al.[8] (1995)	73	45%	Not mentioned		X
Ratnaraj et al.[9] (2002)	50	52%	44%		X
Apfelbaum et al.[10] (2000)	900	Not mentioned	3.33%	X	
Mayr et al.[11] (2002)	261	16.1	14.1%	X	
Eleraky et al.[12] (1999)	185	7.5%	2.1%	X	
Bose[13] (1998)	97	5.1	2.1	X	
Robinson et al.[14] (1962)	56	3.5	7.1	X	
Johnston and Crockard[15] (1995)	50	12%	Not mentioned	X	
Lunsford et al.[16] (1980)	253	5%	3%	X	
Grisoli et al.[17] (1989)	122	Not mentioned	1	X	
Wilson and Campbell[18] (1977)	71	3%	3%	X	

The discrepancy between surgeon notes and patient surveys with regard to postoperative outcomes has been suggested previously. Lieberman et al.[1] found substantial differences between unblinded surgeons and patients with regard to their assessments of the outcome of total hip arthroplasty, with the patient assessments being worse. Heary et al.[2] found that the prevalence of postoperative pain at the iliac crest donor site as reported on patient surveys was significantly higher than that documented in surgeon records (34% compared with 8%, $p < 0.0001$). While that study highlighted the important differences in data sources, its applicability was limited because it was based on the experience of a single surgeon and was performed at a time remote from surgery, with only 52% of consecutive patients participating.

The current study is strengthened by its multi-institutional design, the fact that the surgeons were blinded, and the high percentage of consecutive patient involvement; however, it also has multiple limitations. First, dysphonia and dysphagia were selected as representative subjective adverse outcomes after cervical spine surgery. It is certainly possible that the participating surgeons may have been more or less likely to record the presence of other adverse postoperative symptoms such as discomfort at the iliac crest donor site, axial discomfort, or incisional neuroma dysesthesias. However, our finding that all four surgeons' records consistently underreported the two symptoms suggests that the phenomenon of underreporting of adverse symptoms in surgeon records is commonplace. The results of the present study, considered along with the results of the study by Heary et al.[2], strongly suggest that substantial underreporting of adverse outcomes in surgeon records may not be limited to an isolated few symptoms. The representative nature of the surgeon records utilized in this study is further substantiated by the similarity of the prevalence of symptoms as documented in these records (5% to 11%) with that in other published studies involving the use of surgeon records (Table IV).

A second limitation of the study is that the presence of dysphonia or dysphagia was determined on the basis of patient reports rather than according to specific scientific criteria. Although the symptoms were not verified, they were important enough, from the patients' perspective, to be reported when the patients were asked about them. In addition, they were described as moderate or severe by 63% of the patients and, in many cases, they had persisted for six months or more. A patient's report of subjective symptoms is an important element in the establishment of a diagnostic and treatment pathway. In a similar manner, subjective reports of adverse symptoms should be considered an important measure of a treatment's success.

A third limitation is that the delay between the patient survey and the various follow-up visits may have introduced the potential for recall bias on the part of the patients. The time delay between the patients' most recent office visit and the survey was typically one to three months. Previous studies have demonstrated that recall is typically influenced negatively by a prolonged interval between questioning[3-6]. Stated another way, patients are less likely to report the presence of symptoms with the passage of time. If the patients in the current study indeed erred by underreporting the presence of symptoms, then the discrepancy between the surgeon records and the patient surveys may have been even greater than reported.

A fourth limitation is that we did not conduct this study prospectively. Prospective, concurrent collection of surgeon and patient data was not possible because the participating surgeons were to remain uninformed of the study design. Unfortunately, informing surgeons of the study design likely would have altered their emphasis on the measured symptoms.

Our results suggest that the prevalence of adverse outcomes may be seriously underreported in studies that rely on the retrospective analysis of office notes. Ideally, investigators should avoid the use of surgeon records as a source of data on subjective adverse outcomes and instead should utilize symptom-specific patient surveys or prospective independent data-collection methods. On the basis of these findings, we recommend that conclusions drawn from clinical studies that employ physicians' or surgeons' narrative records as a data source be tempered in their scope unless they are supported by independent patient-derived data or other objective sources. ■

Charles C. Edwards II, MD
Yekaterina Karpitskaya, MD
Carl Lauryssen, MD
K. Daniel Riew, MD
Department of Orthopaedic Surgery (C.C.E. II, Y.K., and K.D.R.) and Department of Neurological Surgery (C.L.), Washington University School of Medicine, One Barnes-Jewish Hospital Plaza, Suite 11300, West Pavilion, St. Louis, MO 63110. E-mail address for K.D. Riew: riewd@msnotes.wustl.edu

Chuck Cha, MD
John G. Heller, MD
S. Timothy Yoon, MD
Department of Orthopaedic Surgery, Emory University School of Medicine, 2165 North Decatur Road, Atlanta, GA 30033

The authors did not receive grants or outside funding in support of their research or preparation of this manuscript. They did not receive payments or other benefits or a commitment or agreement to provide such benefits from a commercial entity. No commercial entity paid or directed, or agreed to pay or direct, any benefits to any research fund, foundation, educational institution, or other charitable or nonprofit organization with which the authors are affiliated or associated.

References

1. **Lieberman JR, Dorey F, Shekelle P, Schumacher L, Thomas BJ, Kilgus DJ, Finerman GA.** Differences between patients' and physicians' evaluations of outcome after total hip arthroplasty. *J Bone Joint Surg Am.* 1996;78:835-8.
2. **Heary RF, Schlenk RP, Sacchieri TA, Barone D, Brotea C.** Persistent iliac crest donor site pain: independent outcome assessment. *Neurosurgery.* 2002;50:510-7.
3. **Boyer GS, Templin DW, Goring WP, Cornoni-Huntley JC, Everett DF, Lawrence RC, Heyse SP, Bowler A.** Discrepancies between patient recall

and the medical record. Potential impact on diagnosis and clinical assessment of chronic disease. *Arch Intern Med.* 1995;155:1868-72.

4. **Carey TS, Garrett J, Jackman A, Sanders L, Kalsbeek W.** Reporting of acute low back pain in a telephone interview. Identification of potential biases. *Spine.* 1995;20:787-90.
5. **Epstein NE, Hood DC.** A comparison of surgeon's assessment to patient's self analysis (short form 36) after far lateral lumbar disc surgery. An outcome study. *Spine.* 1997;22:2422-8. Erratum in: *Spine* 1998;23:284.
6. **Dawson EG, Kanim LE, Sra P, Dorey FJ, Goldstein TB, Delamarter RB, Sandhu HS.** Low back pain recollection versus concurrent accounts: outcomes analysis. *Spine.* 2002;27;984-94.
7. **Winslow CP, Winslow TJ, Wax MK.** Dysphonia and dysphagia following the anterior approach to the cervical spine. *Arch Otolaryngol Head Neck Surg.* 2001;127:51-5.
8. **Stewart M, Johnston RA, Stewart I, Wilson JA.** Swallowing performance following anterior cervical spine surgery. *Br J Neurosurg.* 1995;9:605-9.
9. **Ratnaraj J, Todorov A, McHugh T, Cheng MA, Lauryssen C.** Effects of decreasing endotracheal tube cuff pressures during neck retraction for anterior cervical spine surgery. *J Neurosurg.* 2002;97(2 Suppl):176-9.
10. **Apfelbaum RI, Kriskovich MD, Haller JR.** On the incidence, cause, and prevention of recurrent laryngeal nerve palsies during anterior cervical spine surgery. *Spine.* 2000;25:2906-12.
11. **Mayr MT, Subach BR, Comey CH, Rodts GE, Haid RW Jr.** Cervical spinal stenosis: outcome after anterior corpectomy, allograft reconstruction, and instrumentation. *J Neurosurg.* 2002;96(1 Suppl):10-6.
12. **Eleraky MA, Llanos C, Sonntag VK.** Cervical corpectomy: report of 185 cases and review of the literature. *J Neurosurg*. 1999;90(1 Suppl):35-41.
13. **Bose B.** Anterior cervical fusion using caspar plating: analysis of results and review of the literature. *Surg Neurol.* 1998;49:25-31.
14. **Robinson RA, Walker AE, Ferlic DC, Wiecking DK.** The results of anterior interbody fusion of the cervical spine. *J Bone Joint Surg Am.* 1962;44:1569-87.
15. **Johnston FG, Crockard HA.** One-stage internal fixation and anterior fusion in complex cervical spinal disorders. *J Neurosurg.* 1995;82:234-8.
16. **Lunsford LD, Bissonette DJ, Jannetta PJ, Sheptak PE, Zorub DS.** Anterior surgery for cervical disc disease. Part 1: Treatment of lateral cervical disc herniation in 253 cases. *J Neurosurg.* 1980:53;1-11.
17. **Grisoli F, Graziani N, Fabrizi AP, Peragut JC, Vincentelli F, Diaz-Vasquez P.** Anterior discectomy without fusion for treatment of cervical lateral soft disc extrusion: a follow-up of 120 cases. *Neurosurgery.* 1989;24:853-9.
18. **Wilson DH, Campbell DD.** Anterior cervical discectomy without bone graft. Report of 71 cases. *J Neurosurg.* 1977;47:551-5.