Evaluation of results in lumbar spine surgery

Alf L Nachemson

Department of Orthopedics, University of Gothenburg, Sahlgren Hospital, S-413 45 Gothenburg, Sweden
Tel +46-31 601815. Fax +46-31 602630

It is likely that in most industrialized countries today too much lumbar spine surgery is performed. This is particularly true in the North American continent and has been noted also in the lay press (U.S. News and World Report, January 30, 1989).

Everything we treat must be evaluated against the following basic questions:

1) What causes problem?
2) What is the natural history?
3) Is our diagnosis well defined and our test accurate?
4) Is our treatment proven effective?

The Quebec-report (Nachemson et al. 1987) clearly stated that "low back pain without an anatomic disorder objectively demonstrated is not an indication for spinal surgery and that surgery, including chemonucleolysis, is indicated only for disc herniations after conservative treatments have failed" Most studies indicate that this should be allowed for at least six weeks after the onset of the nerve root pain.

It was also concluded that "a second spinal intervention is indicated only in exceptional circumstances."

Diagnosis

Radiologic studies are with the few exceptions of infections, tumors etc. of poor diagnostic validity in patients with low back pain. One example is the diagnosis of lumbar disc degeneration which, however, is believed by many physicians to cause the pain. Patients are labeled "degenerative" which runs the risk of making them further ill.

Diagnostic labels without scientific proof of validity for low back pain (Nachemson 1992) include the facet syndrome, isolated disc resorption and segmental instability.

Abnormal diagnostic behavior leads to abnormal illness behavior which in its turn leads to abnormal treatment behavior, also by the part of the surgeons who not infrequently perform large surgical interventions on patients with ill defined diagnostic labels despite lack of proof of even moderate effectiveness of such operations.

Radiographic measurements for clinical decisions has also been subjective to criticism (Cowell 1990).

The sensitivity, specificity, and validity of the various diagnostic procedures must be further scrutinized. Only in patients with disc hernia and spinal stenosis has this been done and clearly demonstrated the fairly poor specificity of water soluble myelogram, CT, and MRI, although they are 95% sensitive.

Another such test that still awaits its validation is discography. The gold standard here would be an operative procedure with satisfactory outcome based on pain reproduction on discography.

Much quoted studies supporting discography (Colhoun et al. 1988, Walsh et al. 1990) have not, admittedly, proven the specificity, sensitivity, and validity. Walsh et al. (1990) found that pain behavior was also highly significant in the responders, disregarding the fact that Waddell in several articles during the eighties has demonstrated that a high level of pain or illness behavior precludes a satisfactory result from low back pain surgery (Waddell et al. 1980, Waddell et al. 1982, Waddell and Main 1984, Waddell et al. 1984, Waddell 1992).

In one study Esses et al. (1989) found the predictive value of the pain reproduction discography non-valid for a least short term results of lumbar spine fusion.

In a recent study (Hess et al. 1992) where the outcome of 93 patients, 236 injected discs, with a solid fusion, followed for two years by an unbiased observer, found that pain response preoperatively was a poor predictor of clinical outcome. These authors clearly stated that pain response should not be relied upon diagnostically for selection of fusion levels.

As previously mentioned in the section on disc hernia surgery there has been an increase in these operations in some parts of the world and also an increase in lumbar spine fusion has been noted (Deyo et al. 1992, Turner et al. 1992). It could be that the semi-invasive and perhaps "simpler" methods, like injections or automated percutaneous discectomy, increase the affinity for doing something in patients not following the criteria mentioned previously (Nachemson 1993).

Lumbar spine fusion

From an early study (Flynn and Hoque 1979) it was concluded that when fusion was performed for "degenerative disc disease" those patients who were on workers compensation had only a 20% good result, while when fusion was performed for spondylolisthesis there were similar good result, 85–90%, in the private and workers comp-cases. This may suggest that when we operate on "motion segment instability" we may get better results. Certainly most spine surgeons agree that motion segment instability exist as a clinical problem but so far is poorly defined (Stokes and Frymoyer 1987). Attempts to further our diagnostic capabilities in this area have been made (Friberg 1987, Selvik 1989, Johnsson et al. 1990).

With one notable exception there have been no prospective randomized trials on patient outcomes after fusion (Herkowitz and Kurz 1991, Turner et al. 1992).

Numerous retrospective follow-up studies to be found in Deyo's and Turner's articles (1992) have failed to substantiate that the solid fusion really is substantially better than when the attempted fusion resulted in pseudarthrosis. This finding alone is remarkable and actually should raise serious questions about the whole operation. There are certain kinds of fusion procedures like the PLIF that also have a rather high complication rate, while the suggested advantages like increase in disc heights following the surgery has been refuted by Dennis et al (Dennis et al 1989).

With many studies showing a high rate of pseudarthrosis, the advent of pedicle fixation has caused surgeons and instrument makers to suggest various types of internal fixation procedures. Their efficacy has not, however, been demonstrated in clinical studies. Retrospective studies of large materials have reported high complication rates: Scoliosis Research Society reported 26%, Esses et al. 27% (Lowe 1988, Esses and Sachs 1991). Without the support of studies that definitely show an advantage of these new internal fixation methods, it should be of some concern for the surgeons and the patients alike. The latter should be informed about the dangers and non-proven value for his condition. Physicians should be aware of the Hippocratic oath they have taken; *Primum non nocere.*

Some retrospective studies (Zucherman et al. 1992) has also cast some doubts on the superiority of metallic internal fixation and in a recent report (McGuire et al. 1992) a prospective randomized study with in situ fusion for low grade spondylolisthesis no benefit could be demonstrated for the VSP-plates. The study reported on 50 patients, minimum two year follow-up by an unbiased observer and found no difference in fusion rate, or clinical benefit for the patients. It may be time to warn the public that "too much treatment may be dangerous for your health."

The multiply operated back

Only when a new hernia is found the results from repeat surgery seem to be worthwhile. If no new hernia is found at exploration, the collective results from the literature show only a 30% relief of the back and leg pain.

Repeat surgery is performed in the U.S. in 15% of the cases primarily operated for a lumbar disc disease, while the comparative figure in Europe is 6-7%. Long et al. found in an article on the clinical features of the failed back syndrome that 66% of the patients originally before their first operation did not meet the afore mentioned A.A.O.S. criteria (Long et al. 1988, Nachemson 1993).

The diagnosis of a recurrent true disc herniation is difficult but as we heard in this meeting (Holtås 1993), gadolinium enhanced magnetic resonance imaging seem to be promising. A clearly pain free interval lasting more than 12 months after the previous surgery and then a renewed attack of sciatica seems to speak in favor of a new disc hernia, while if the pain relief was only 6 months or less other causes like scar tissue or arachnoiditis or internal nerve root injury could be the reason for the recurrent pain.

Also, the pain description should be distinctive, not diffuse, and intermitting rather than continuous.

Waddell has demonstrated that our failed back patients have much more pain, disability, pain drawing magnification and inappropriate Waddell signs, than for example our failed total hips (Waddell 1992). In a study with a long follow up by an unbiased observer North et al. found that only 1/3 of 102 patients reoperated had a successful result which was said to be achieved when there was a modest 50% or more of reduction of pain and subjective satisfaction (North et al. 1991). That repeat surgery is difficult and not so rewarding was also underlined by Lauerman et al. (Lauerman et al. 1992).

Why then are surgeons still prone both to use internal fixation devices in large amounts and to reoperate again and again. Part of this could be blamed on retrospective series of low quality that have been published in the medical literature. The responsibility here rests heavily on the editors of scientific medical journals (Nachemson and LaRocca 1987, Goodfellow 1992).

It is quite clear that only prospective randomized trials are conclusive when it comes to evaluate treatment effectiveness (Rudicel and Esdaile 1985, Bloch 1987, Gartland 1988). Adequate prospective cohort studies are also few but could at least tell us something, while the ordinary retrospective studies have a

low level of significance for effectiveness.

In the retrospective studies, however, it is absolutely necessary that

1) The patient selection is clearly given. There should be no mixture of patients with different conditions and the number of patients in each distinctively described category must exceed at least thirty.

2) There should be a proper objective evaluation before as well as at follow-up using various now well established baseline and outcome scales.

3) The follow-up must be personal and should not be done by questionnaires or telephone interviews. They should be performed by a non-treating observer, i.e. never by the surgeon himself. That this is important has been demonstrated (Ejeskär et al. 1983, Roberts et al. 1984). Also, each case should have a minimum follow-up of two years and the number of subjects followed must exceed 80% of included consecutive original patients. The patients not retrieved should be evaluated for baseline characteristics and be demonstrated not to deviate from the important clinical and demographic variables of the remaining group. If these prerequisites are not followed, (and they can easily be found in the Materials and Methods section of any article) then the article should not be published, and if it has been done so in the past it should not be read or quoted. The importance of these objectives has been the subject of numerous editorials in various papers (Bloch 1987, Nachemson & LaRocca 1987, Goodfellow 1992).

The importance of psychologic factors

This decade has been declared "The decade of the brain" which should be noted also by spine surgeons, for whom psychologic knowledge is more important than knowledge of biomechanics.

We often misinterpret suffering and ineffective coping in patients with chronic low back pain. We are too hung up on the presumed physical problem which in several studies has been shown to be irrelevant, especially in chronic low back pain and sciatica patients. We should make use of a pain drawing, of Waddell illness behavior signs, his subjective disability questionnaire, the UAB behavior scale and other such validated measures of both health and pain beliefs, of fear and distress expressing themselves as pain or illness behavior.

In cases demonstrating high scores on any of these scales surgery must be avoided unless a definite validated diagnostic condition, necessitating an intervention, has been demonstrated.

Conclusions

In my view, with few exceptions, further surgical interventions in the spine should be stopped unless they are part of some type of scientific trial, be it a multicenter or a prospective controlled trial. The exceptions include disc hernia, recurrent disc hernia with proper preoperative conservative treatment, clear-cut spinal or root canal stenosis, tumors, fractures, and infections. For instability demonstrated by large movements exceeding 6-7 mm, in some spondylolisthesis patients and after previous laminectomy, there is probably also an indication.

In summary I hope that I have made it clear why we need more properly controlled randomized and prospective studies.

References

Bloch R. Methodology of clinical back pain trials. Spine 1987; 12 (5): 430–2.

Colhoun E, McCall I W, Williams L, Cassar Pullicino V N. Provocation discography as a guide to planning operations on the spine. J Bone Joint Surg (Br) 1988; 70 (2): 267–71.

Cowell H R. Editorial. Radiographic measurements and clinical decisians. J Bone Joint Surg (Am) 1990; 72 (3): 319.

Dennis S, Watkins R, Landaker S, Dillin W, Springer D. Comparison of disc space heights after anterior lumbar interbody fusion. Spine 1989; 14 (8): 876–8.

Deyo R A, Cherkin D C, Loeser J D, Bigos S J, Ciol M A. Morbidity and mortality in association with operations on the lumbar spine. The influcence of age, diagnosis and procedure. J Bone Joint Surg (Am) 1992; 74 (4): 536–43.

Ejeskär A, Nachemson A, Herberts P, Lysell E, Anderson G, Irstam L, Peterson L E. Surgery versus chemonucleolysis for herniated lumbar discs. A prospective study with random assignment. Clin Orthop 1983; 174: 236–42.

Esses S I, Botsford D J, Kostuik J P. The role of external spinal skeletal fixation in the assessment of low back disorders. Spine 1989; 14 (6): 594–601.

Esses S I, Sachs B L. Complications of pedicle screw fixation. In: Proceedings of the Twenty-fifth Anniversary Meeting of Scoliosis Research Society, Minneapolis, Minnesota, USA 1991: 185.

Flynn J C, Hoque M A. Anterior fusion of the lumbar spine. End-result study with long-term follow-up. J Bone Joint Surg (Am) 1979; 61 (8): 1143–50.

Friberg O. Lumbar instability: A dynamic approach by traction–compression radiography. Spine 1987; 12 (2): 119–29.

Gartland J J. Orthopaedic clinical research. Deficiencies in experimental design and determination of outcome. J Bone Joint Surg (Am) 1988; 70 (9): 1357–63.

Goodfellow J. Malignancy and joint replacement: Editorial: J Bone Joint Surg (Br) 1992; 74 (5): 645.

Herkowitz H N, Kurz L T. Degenerative lumbar spondylolisthesis with spinal stenosis. A prospective study comparing decompression with decompression and intertransverse process arthrodesis. J Bone Joint Surg (Am) 1991; 73 (6): 802–8.

Hess W F, Jackson R P, Ebelke D K, Crawley S A, Arnett K W. Pain response by discography as a predictor of clinical outcome in patients with solid posterolateral lumbosacral fusion. In: Proceedings of the Twenty-seventh Annual Meeting of Scoliosis Research Society, Kansas City, Missouri, USA 1992: 82.

Holtås S. The degenerative lumbar spine—radiological aspects. Acta Orthop Scand 1993: submitted.

Johnsson R, Selvik G, Strömqvist B, Sundén G. Mobility of the lower lumbar spine after posterolateral fusion determined by roentgen stereophotogrammetric analysis. Spine 1990; 15 (5): 347–50.

Lauerman W C, Bradford D S, Ogilvie J W, Transfeldt E E. Results of lumbar pseudarthrosis repair. J Spinal Disord 1992; 5 (2): 149–57.

Long D M, Filtzer D L, BenDebba M, Hendler N H. Clinical features of the failed back syndrome. J Neurosurg 1988; 69 (1): 61–71.

Lowe T. Morbidity and mortality report. Read at the 23rd Annual Meeting of Scoliosis Research Society, Baltimore, USA 1988.

McGuire R A, Amundson G M, Jackson M S. The use of primary internal fixation in spondylolisthesis. In: Proceedings of the Twenty-seventh Annual Meeting of Scoliosis Research Society, Kansas City, Missouri, USA 1992: 115.

Nachemson A with Spitzer W O et al. Scientific approach to the assessment and management of activity-related spinal disorders. A monograph for clinicians. Report of the Quebec Task Force on Spinal Disorders. Spine (Suppl 7S) 1987; S1–S59.

Nachemson A L, LaRocca H. Editorial: Spine 1987. Spine 1987; 12 (5): 427–9.

Nachemson A L. Newest knowledge of low back pain. A critical look. Clin Orthop 1992; 279: 8–20.

Nachemson A L. Lumbar disc herniation - conclusions. Acta Orthop Scand 1993; submitted.

North R B, Campbell J N, James C S, Conover-Walker M K, Wang H, Piantadosi S, Rybock J D, Long D M. Failed back surgery syndrome: 5-year follow-up in 102 patients undergoing repeated operation. Neurosurgery 1991; 28 (5): 685–91.

Roberts N, Smith R, Bennett S, Cape J, Norton R, Kilburn P. Health beliefs and rehabilitation after lumbar disc surgery. J Psychosom Res 1984; 28 (2): 139–44.

Rudicel S, Esdaile J. The randomized clinical trial in orthopaedics: Obligation or option? J Bone Joint Surg (Am) 1985; 67 (8): 1284–93.

Selvik G. Roentgen stereophotogrammetry. A method for the study of the kinematics of the skeletal system. Acta Orthop Scand (Suppl 232) 1989; 60: 1–51.

Stokes I A F, Frymoyer J W. Segmental motion and instability. Spine 1987; 12 (7): 688–91.

Turner J A, Ersek M, Herron L, Deyo R. Surgery for lumbar spinal stenosis: Attempted meta-analysis of the literature. Spine 1992; 17 (1): 1–8.

Turner J A, Ersek M, Herron L, Haselkorn J, Kent D, Ciol M A, Deyo R. Patient outcomes after lumbar spinal fusions. JAMA 1992; 268 (7): 907–11.

Waddell G, McCulloch J A, Kummel E, Venner R M. Nonorganic physical signs in low back pain. 1979 Volvo Award in clinical science. Spine 1980; 5 (2): 117–23.

Waddell G, Main C J, Morris E W, Venner R M, Rae P S, Sharmy S H, Galloway H. Normality and reliability in the clinical assessment of backache. Br Med J 1982; 284 (6328): 1519–23.

Waddell G, Main C J. Assessment of severity in low back disorders. Spine 1984; 9 (2): 204–8.

Waddell G, Main C J, Morris E W, DiPaola M, Gray I C M. Chronic low back pain, psychological distress and illness behavior. Spine 1984; 9 (2): 209–13.

Waddell G. Understanding the patient with backache. In: The lumbar spine and back pain (Ed. Jayson M I V). Churchill Livingstone 1992; 25: 469–85.

Walsh T R, Weinstein J N, Spratt K F, Lehmann T R, Aprill C, Sayre H. J Bone Joint Surg (Am) 1990; 72 (7): 1081–8.

Zucherman J, Hsu K, Picetti G, White A, Wynne G, Taylor L. Clinical efficacy of spinal instrumentation in lumbar degenerative disc disease. Spine 1992; 17 (7): 834–7.