

Results of treatment of lumbar spine disorders

Effects of assessment techniques and confounding factors

Charles G Greenough

Department of Orthopedics, Middlesbrough General Hospital, Ayresome Green Lane, Middlesbrough, Cleveland TS5 5AZ, United Kingdom. Tel +642-850222. Fax + 642-850710.

The assessment of the results of spinal surgery has been complicated by the large number of patient and assessment related factors which may have a considerable effect on the outcome. Comparison of fourteen different outcome measures obtained from published reports in a group of over two hundred patients who had undergone lumbar intervertebral disc excision more than ten years previously revealed that the proportion of "successes" was significantly influenced by the outcome measure used, ranging from 97% to 60% (Howe and Frymoyer 1985). Subjective measures, for example the patients opinion of the operation, gave a higher proportion of "successful" results when compared with objective measures such as return to original employment.

When considering outcome assessment it is important to distinguish between two quite different results, the technical outcome and the clinical outcome. Technical outcome is the answer to the question "Has the surgeon achieved what he set out to achieve?" Clinical outcome is the answer to the question "Has the patient benefited from the procedure?" A successful technical result does not necessarily mean a successful clinical result and visa-versa. It is wrong, therefore, to include measurements of technical success in composite outcome scores.

Examples of technical outcome are the restoration of dural pulsation in spinal decompression, the reduction of a spondylolisthesis or the achievement of bony union in spinal fusion. The measurement of technical outcome is itself not easy, for example in spinal fusion. The most important feature according to Stauffer and Coventry (1972) was the demonstration of trabeculae crossing the fusion site. This method almost certainly over estimates the union rate. Percy and Borough (1982) using an accurate biplanar radiological technique in flexion and extension demonstrated significant motion between the vertebrae in 6 out of 10 patients who had conventional evidence of union with the appearance of trabeculae crossing the fusion site. Further, Laasonen and Soini (1989) found in a group

of 25 patients with nonunion demonstrated on CT scanning that assessment using plain films and flexion/extension views falsely indicated union in 16 cases.

Clinical outcome has been measured in many ways, but the most important are clinical examination (or physical impairment), the patient's subjective opinion of the operation and disability scores of various sorts. Physical examination suffers from a number of significant difficulties. Some attributes are affected by psychological disturbance (Rae et al 1984). Some physical measurements display a considerable diurnal variation (Porter and Trailescu 1990). Interobserver error is a problem (McCombe et al 1989). Intraobserver error is also possible over the long period of follow up required. Observer bias will exist.

The patient's subjective opinion will vary depending on questioner. There is no agreed format for a questionnaire of patient opinion. What questions should be asked? An agreed format would at least allow comparison of the results one study with another.

Turning to disability measures, a number of these have been published over the years. Most are in the form of a patient self-report questionnaires and there have been few studies directly comparing one scale with another. One such study compared the Oswestry Disability Index, the Disability score of Waddell and the Low-Back Outcome Score (Greenough and Fraser 1992). The Low-Back Outcome Score uses 13 different items, totalling 75 points, and the items are weighted for pain and the energy involved (Figure 1). The three scores were assessed in a group of patients followed up after conservative therapy for low-back pain. The Low-Back Outcome Score proved to be more discriminating, particularly in the patients with better results. The least disabled 20% of the studied population all scored zero on the Waddell Disability Score. They were split between 3 points on the Oswestry Disability Index but between 4 points on the Low-Back Outcome Score. The Low-Back Outcome

Figure 1. The Low-Back Outcome Score as given to the patient. Scores are omitted, scoring for the pain scale: 7-10=0; 5-6=3; 3-4=6; 0-2=9

Back Pain Clinic				
Please mark on the line below how much pain you have had from your back on average over the past week:				
Please tick the answer which most closely describes you on each of the following six sections:				
Do you have to rest during the day because of pain		At present, are you working		
Not at all	(6)	Full time at your usual job	(9)	
A little	(4)	Full time at a lighter job	(6)	
Half the day	(2)	Part time	(3)	
Over half the day	(0)	Not working	(0)	
How often do you have a consultation with a doctor or have any treatment (e.g. physiotherapy) for your pain?		At present, can you undertake sports or active pursuits (e.g. dancing)		
Never	(6)	As much as usual	(9)	
Rarely	(4)	Almost as much as usual	(6)	
About once a month	(2)	Some, much less than usual	(3)	
More than once a month	(0)	Not at all	(0)	
How often do you have to take pain killers for your pain?		At present, can you undertake household chores or odd jobs		
Never	(6)	Normally	(9)	
Occasionally	(4)	As much as usual but slowly	(6)	
Almost every day	(2)	A few, not as many as usual	(3)	
Several times each day	(0)	Not at all	(0)	
Please tick the box that describes best how much your back pain affects each of the following six activities:				
	No effect	Mildly / Not much	Moderately / Difficult	Severely / Impossible
Sex life	(6)	(4)	(2)	(0)
Sleeping	(3)	(2)	(1)	(0)
Walking	(3)	(2)	(1)	(0)
Sitting	(3)	(2)	(1)	(0)
Travelling	(3)	(2)	(1)	(0)
Dressing	(3)	(2)	(1)	(0)

Score was also the most comprehensive in explaining the variance of the outcome of the patients.

On top of these difficulties in patient assessment there are a number of confounding factors that are related to the patients themselves. Compensation, psychological distress, age, sex, and social group are most important in this regard (Greenough and Fraser 1989). Data from two studies will be presented to illustrate the important effects of these confounding factors, one is a study of conservative treatment and the other the results of anterior interbody lumbar spinal fusion.

In the first study, 300 patients were assessed after a course of conservative treatment for low-back pain including physiotherapy, epidural injections, facet joint injections, education classes and exercises. All the patients related the onset of their back pain to an identifiable incident, had never had a fracture or dislocation of the spine and none of them had any surgery during the period of the study. They were all of work-

ing age at the time of the injury. Most of them (45%) were diagnosed as mechanical (discogenic) pain. Pro-lapsed intervertebral discs as proven by CT scanning were present in 15 percent, facet joint arthropathy in 15 percent, spondylolysis or spondylolisthesis in 7 percent and soft tissue strain in 10 percent. They were reviewed at an median follow up of 51 months and a 91 percent follow-up was achieved. At review the Low-Back Outcome Score was administered and the patients were seen by an independent observer. The results were analyzed using multiple linear regression analysis using the Outcome Score as the dependent variable and 16 patient and treatment related factors as the independent variables. The results are expressed here as (coefficient, *p*-value).

The patient's sex was important with men faring somewhat better (5.5 ± 12 , $p < 0.005$). Increasing age at injury reduced the results (-0.3 ± 0.09 , $p < 0.005$). Increasing time off work reduced the outcome

Table 1. Subjective opinions at review, number (%)

	Compensation	Noncompensation	All patients
"Almost complete relief"	15 (18)	15 (38)	30 (24)
"Good deal of relief"	38 (45)	16 (41)	54 (44)
"Only a little relief"	23 (27)	4 (10)	27 (22)
"No relief or worse"	9 (10)	4 (10)	13 (10)

$p < 0.05$, Chi-square.

Table 2. Low back Outcome scores at review, number (%)

	Compensation	Noncompensation	All patients
Excellent (65-75)	8 (10)	13 (43)	21 (17)
Good (50-64)	19 (22)	10 (25)	29 (23)
Fair (30-49)	34 (40)	10 (25)	44 (35)
Poor (0-29)	24 (28)	7 (17)	31 (25)

$p < 0.01$, Chi-square.

(-0.17 ± 0.06 , $p < 0.005$) and patients in lower socio-economic groups did worse (-2.0 ± 0.9 , $p < 0.05$).

Conversely, neurological deficits, either loss of a reflex or a specific weakness of a specific muscle group, had no influence on the outcome ($p < 0.22$). Diagnosis made no significant difference to the outcome ($p < 0.20$). The total variance explained by the diagnosis was only 2.6 percent. Of the treatments, only educational classes could be shown to have any significant effect whatever on the outcome and that was only statistically significant in a small subset, the mechanical back pains in the compensation group ($p < 0.05$).

Of far greater importance were compensation and psychological disturbance at presentation. Compensation greatly reduced the outcome (-6.1 ± 1.3 , $p < 0.0001$) as did psychological disturbance (-7.9 ± 2.5 , $p < 0.005$). In this study the total variance of the outcome explained by all of the factors examined was 47 percent. Compensation and psychological disturbance between them accounted for 35 percent of the explained variance. Thus in this study, patient related confounding factors were of far greater significance than the diagnosis, physical findings or treatment.

The second study relates to the results of anterior interbody lumbar spinal fusion. 67 men and 69 women underwent surgery, of whom 93 were claiming compensation. The median age at operation was 41 (17-62) years and the median follow-up was 40 (24-82) months. The patients were reviewed by an independent observer and the patients were assessed using the Low-Back Outcome Score. A 90-percent follow-up was achieved.

Turning to the technical results, fusion was obtained in 100 patients (74%). Interestingly, compensation had a major influence on the technical results, with 61 of 93 compensation patients fusing compared to 39 of 43 noncompensation patients ($p < 0.01$). No differences were observed, however, between the fused patients and the nonfused patients in the Outcome Score (46 [1-75] vs. 43 [19-75], n.s.), the patient's subjective opinion on a 4-point scale (2 [0-3] vs. 2 [0-3], n.s.) or the patient's reported pain (4 [0-10] vs. 5 [1-8], n.s.); as there were so few nonunions in the noncompensation patients only the compensation patients are reported here. Thus in this study the technical result of bony union was irrelevant to the clinical results and obviously needs to be reported separately.

Compensation made a very large difference to the results both as expressed by the patient's subjective opinion (Table 1) and by the Low-Back Outcome Score (Table 2). Although 50 percent of patients unemployed preoperatively were able to return to work, only 42 of 92 compensation patients were employed compared to 35 of 43 noncompensation patients ($p < 0.001$).

Psychological disturbance at review as measured by the Zung depression scale and the Modified Somatic Perception Questionnaire (Greenough and Fraser 1991) has an important influence, particularly in the compensation patients. In this group disturbed patients had poorer results as measured by the Outcome Score (28 [3-55] vs. 50 [1-75], $p < 0.0001$), patient satisfaction (1 [0-3] vs. 2 [0-3], $p < 0.0001$) and patient pain rating (6 [2-10] vs. 4 [0-10], $p < 0.0001$). Using multiple linear regression analysis, the important prog-

nostic factors for a poor result were compensation (-6.1 , $p < 0.05$), reoperation (-8.5 , $p < 0.05$) and psychological disturbance at review (-0.56 , $p < 0.00001$).

It is of interest that using the patient's subjective assessment 67 percent of the patients were either very satisfied or had a good deal of relief. Using the Low-Back Outcome Score, however, only 40 percent fall into the excellent or good categories. The Low-Back Outcome Score was designed to be also used in non-surgically treated patients who may well be less disabled at presentation. It is important therefore to score the patients before the intervention in order to adequately define the study population.

In summary, the patient population and initial disability, the assessment technique, and confounding factors all have a major role in determining the "results" of a treatment. Indeed, these factors may explain more of the variance of the outcome than the condition or treatment itself. In future studies a number of points must be considered:

1. The population must be adequately described prior to the intervention.
2. Technical results should be described separately.
3. Confounding factors must be explicitly described and included in the analysis.
4. Subjective results alone are unsatisfactory and a standard format should be adopted.
5. Outcome must be measured using a recognized instrument which is discriminating and comprehensive.
6. The same outcome instrument should be applied both before (initial severity) and after intervention.

References

- Greenough C G, Fraser R D. The effects of compensation on recovery from low-back injury. *Spine* 1989; 14: 947-955
- Greenough C G, Fraser R D. Comparison of eight psychometric instruments in unselected patients with back pain. *Spine* 1991; 16: 1068-1074.
- Greenough C G, Fraser R D. Assessment of outcome in patients with low-back pain. *Spine* 1992; 17: 36-41.
- Howe J, Frymoyer J W. The Effects of questionnaire design on the determination of end results in lumbar spine surgery. *Spine* 1985; 10: 804-805.
- Laasonen E M, Soini J. Low back pain after lumbar fusion. Surgical and computed tomographic analysis. *Spine* 1989; 14: 210-213.
- McCombe P F, Fairbank J C T, Cockersole B C, Pynsent P B. Reproducibility of physical signs in low-back pain. *Spine* 1989; 14: 908-918.
- Pearcy M, Burrough S. Assessment of bony union after interbody fusion of the lumbar spine using a biplanar radiographic technique. *J Bone Joint Surg (Br)* 1982; 64B: 228-232.
- Porter R W, Trailescu I F. Diurnal changes in straight leg raising. *Spine* 1990; 15: 103-106.
- Rae P S, Waddell G, Venner R M. A Simple technique for Measuring lumbar spinal flexion. *J Roy Coll Surg Edin* 1984; 29: 281-284.
- Stauffer R N, Coventry M B. Posterolateral lumbar-spine fusion. *J Bone Joint Surg (Am)* 1972; 54A: 1195-1204.