Malhar N. Kumar
Frederic Jacquot
Hamilton Hall

Long-term follow-up of functional outcomes and radiographic changes at adjacent levels following lumbar spine fusion for degenerative disc disease

Received: 20 April 2000
Revised: 28 July 2000
Accepted: 9 August 2000
Published online: 9 December 2000
© Springer-Verlag 2000

M. N. Kumar (✉) · H. Hall
Division of Spine Surgery,
Orthopaedic and Arthritic Hospital,
Toronto, Ontario, Canada
e-mail: kumarmd1@yahoo.com

F. Jacquot
Division of Spine Surgery, Institut Calot,
Berck Sur Mer, France

M. N. Kumar
9/1 Sapta Nilaya, 3rd Temple Street,
15th Cross Malleswaram, Bangalore,
India 560003

Abstract There are very few studies with more than 20 years' follow-up of lumbar spine fusions for disc degeneration. Currently, there is a lot of interest in the subject of degenerative changes above the level of fusion; this study is concerned with such changes in the very long term (30 years). Twenty-eight patients showing sound fusion on radiographs following posterior midline spinal fusion performed by a single surgeon between 1968 and 1970 were compared with an age- and gender-matched group of 28 patients who had undergone surgery for degenerative disc disease without fusion during the same period, by the same surgeon and using similar criteria for evaluation (Short Form 36 and Oswestry Disability Index; functional testing using self-paced walk and timed up-and-go; flexion and extension lateral radiographs of the lumbar

spine). In this study, the incidence of radiographic changes at levels above the level of previous involvement was twice as high in the fusion group as in the non-fusion group. However, there was no statistically significant difference between the two groups in the outcomes measured using validated scales and functional testing. The study emphasises the importance of complete evaluation of these patients using validated outcome measurement instruments against the background of radiographic changes and subjective assessment of back pain. It also shows that radiographic changes do not necessarily mean functional impairment in all patients following lumbar spine fusion for degenerative disc disease.

Keywords Lumbar fusion · Adjacent level · Long term

Introduction

Mobile segment changes above the level of fusion have been reported by many authors [6, 8, 11, 12, 13, 17], and the level immediately above the fusion is thought to be at high risk due to the altered biomechanics following fusion. Sagittal plane abnormalities caused by fusion, advanced age of the patient and use of rigid pedicle instrumentation have been held responsible for the involvement of levels above the fusion. Most studies that suggest a high incidence of clinically significant adjacent level degeneration have focused only on symptomatic patients

[11, 22], and have not studied all patients who underwent lumbar fusion. Existing studies that do involve long-term follow-up of patient cohorts following lumbar fusion have certain drawbacks, such as assessment of clinical status using self-reported outcome without validated outcome measurements and performance tests and consideration of radiographic changes alone, without clinical correlation [6, 8, 12, 13, 17]. Interpretation of results from these studies is difficult because of differences in the numbers of patients undergoing the different components of evaluation. Typically, the number of participants decreases at each step, from the time of telephone interview to the time of actual assessment, and the results end up as a mixture of

different strata of data. Inclusion of patients with complications such as obvious pseudarthrosis makes comparison of functional outcome difficult, as it is a complication and not really a long-term outcome. Thus, it is difficult to understand the incidence of clinically significant adjacent level degeneration following lumbar fusion. In this study, efforts have been made to secure a specific control group that resembles the study group very closely, including the length of follow-up, and clinical and radiographic assessments have been made using well-defined and validated criteria. The minimum follow-up in the present study is 30 years, and nearly all patients in the study have the same duration of follow-up.

Materials and methods

Fifty-four consecutive patients underwent posterior midline lumbar fusion (one or two levels) for degenerative disc disease between 1968 and 1970. The minimum follow-up was 30 years. The fusion had been done using an interspinous wiring technique with tibial bone graft. Of the 54 patients originally operated, nine were deceased, five could not be located and four refused to participate in the study. Techniques used to locate the patients included using the original addresses and telephone numbers from the chart at the time of surgery; contacting the family physicians listed in the charts and those who have taken over their practices; medical records departments from the hospitals in the area of the patient's last known address; human resource departments at the place of employment at the time of surgery and CD-ROM North American telephone directories [3].

This left 28 patients available for the study in the fusion group, all of whom underwent clinical and radiographic assessment. A further 28 patients with degenerate disc disease who had been treated with discectomy or decompression without fusion by the same surgeon during the same period were also reviewed. The two groups were age and gender matched. No patient in this series had undergone re-operation for extension of fusion. Degenerative changes following lumbar fusion are time-dependent, and hence a uniform duration of follow-up (rather than the average follow-up) was maintained for all patients.

Outcome evaluation was structured along the following lines:

1. Self-evaluation of the outcome by the patients using validated outcome measurement instruments
2. Clinical evaluation along with functional tests
3. Radiographic evaluation including flexion and extension lateral radiographs of the lumbar spine in the upright position.

Thirty-six patients responded by answering the Short Form 36 (SF36) and Oswestry Disability Index questionnaires. The SF36 [7] and the Oswestry Disability Index [4] are patient-oriented outcome measurement scales that have been shown to possess validity as well as internal consistency for the assessment of low back pain in the context of activities of daily living. Of the 36 patients, five were unable to attend clinical and radiographic assessment. A further three were excluded from the study because they had radiographically obvious pseudarthrosis, as the purpose of the study was to determine the long-term effects of a fused segment of the spine and not the long-term effects of obvious non-union. Patients with spondylolysis with or without spondylolisthesis were not included in the study either, as it is well known that these patients tend to be younger than patients with degenerate disc disease and tend to have better outcomes.

The study subjects were examined clinically by one of the authors and then by a physiotherapist, who performed two functional tests: the self-paced walk and the timed up-and-go.

In the self-paced walk, subjects walk 160 m at their normal pace (neither fast nor slow) without overexerting themselves. The number of paces and the time required for each lap are recorded. In the timed up-and-go test, the subject is seated in a chair and is asked to get up from the chair, stand still momentarily, walk to a line on the floor 3 m away, turn and return to the chair, turn around and sit down. The time taken to complete the test is recorded. After each of these tests, patients were asked to rate their back pain during the test on a visual analogue scale of 0–10 (0=no pain and 10=severe pain). They were also asked to rate their perceived exertion during the test using the Borg rating of perceived exertion. The reliability and reproducibility of these tests in the elderly population has been demonstrated [1, 9, 15].

Radiographs of the lumbosacral spine, including lateral radiographs of the lumbar spine in the upright position during flexion and extension, were obtained. Instability was defined by the criterion of Panjabi and White [24], as sagittal plane translation equal to or more than 25% of the sagittal diameter of the subjacent vertebra. Many studies have used a shift of 3 or 4 mm as the criterion for instability, but this measurement depends on the radiographic magnification and is not reliable unless the latter is kept constant for all patients [6, 8, 12]. Disc space height was measured using the criteria of Pope et al. [18]. Disc space narrowing at levels above the fusion, presence of osteophytic spurs, sclerosis of the vertebral end plates and narrowing of the intervertebral foramina were assessed.

All the radiographs were analysed independently by one of the authors (M.N.K.) and by a consultant radiologist.

For statistical analysis, the *t*-test was used for continuous data and Fisher's exact test was used for assessing the difference between proportions. The chosen level of statistical significance was $P=0.05$.

Results

Table 1 shows the results of SF36 outcome measurement. There was no statistically significant difference between the fusion group and non-fusion group for any of the eight

Table 1 Results of functional outcome using the Short Form 36 (SF36) scale [7]

Parameter	Fusion group Mean (SD)	Non-fusion group Mean (SD)
Physical functioning	57 (31)	61 (29)
Physical role	54 (47)	50 (47)
Bodily pain	64 (29)	59 (28)
General health	65 (28)	60 (28)
Vitality	58 (26)	56 (21)
Social functioning	72 (33)	80 (31)
Emotional role	85 (35)	78 (42)
Mental health	80 (15)	71 (19)

Table 2 Results of self-paced walk test

	Fusion group	Non-fusion group
Time to complete 40 m (mean)	36.6 s	32.5 s
Fast walking speed (mean)	1.26 m/s	1.24 m/s
Reported perceived exertion on the Borg rating scale (mean)	0.95	0.80

Table 3 Results of timed up-and-go test

	Fusion group	Non-fusion group
Time to complete test (mean)	11.8 s	10.2 s
Arm rest use (%)	72.7	73.3
Visual analogue scale (mean)	0.25	0.98
Reported perceived exertion (mean)	0.90	0.66

Table 4 Radiographic results

	Fusion group	Non-fusion group
Significant instability at levels above (%)	14.2	7.4
Osteophytic spurs (%)	42.8	25.9
Disc space narrowing at levels above (%)	35.7	18.5
Mean loss of disc height (mm)	5.7	3.1
Sclerosis in the vertebral end plates (%)	44.3	20.6
Mean narrowing of the vertebral foramina (mm)	3.5	1.9

variables assessed in the SF36 scale. The mean Oswestry Index score in the fusion group was 20.1, while in the non-fusion group it was 23.0. There was no significant difference between the results.

There was also no significant difference between the two groups in the results of the self-paced walk and the timed up-and-go tests (Table 2, Table 3). Radiographic

changes (Table 4) of significant instability (by the White and Panjabi criterion) at levels above the fusion were seen in 14.2% of patients with fusion compared to 7.4% of patients without fusion. Disc space narrowing at levels above was seen in 35.7% of patients with fusion compared to 18.5% of patients without fusion. The difference between the two groups was significant ($P=0.017$).

Discussion

A possible criticism of this study would be that these patients had received a posterior midline fusion, which is not the state-of-the-art today. However, it should be emphasised that the aim of this study is to examine the clinical and radiological changes of a solidly fused segment of the lumbar spine in the long term and to assess the extent of the impact these changes have from a functional viewpoint. It is not the objective of this study to examine the relative merits and demerits of different types of spinal fusion. It is for this reason that we selected only patients with radiographically sound fusion for the study. This would eliminate the effects of pseudarthrosis on the final outcome.

Are we justified in comparing the fusion group to the non-fusion group? The natural history of disc degeneration has been shown to go through the stages of dysfunction, instability and stabilisation along with the increasing age of the patient [10]. Though most degenerated discs go through this process, their clinical presentation varies de-

Table 5 Long-term results of lumbar fusions reported in previous studies (*CT* computed tomography, *MRI* magnetic resonance imaging)

Author	Fusion type	Mean follow-up	Results	Strengths	Drawbacks
Frymoyer et al. [6]	Posterior midline	13.7 yrs	High incidence of degenerative changes in both fusion and non-fusion groups	Large sample size; control group present	High incidence of pseudarthrosis (26%); follow-up is relatively short; no validated outcome scores
Hambly et al. [8]	Posterolateral	22.6 yrs	Anterolisthesis 33%; retrolisthesis 14.2%; dynamic instability 21.4%; disc space narrowing 40.4%; good clinical results 75%	Study examines posterolateral fusion, which is a procedure in current use	Control group is a generic group with 'low back pain'; no validated outcome scores; inclusion of patients with spondylolysis/listhesis; control group did not have the same length of follow-up as the study group
Lehmann et al. [12]	Posterior midline	33 yrs	Instability 45%; stenosis 42%	Long-term follow-up; CT assessment of stenosis	High incidence of pseudarthrosis; no validated outcome scores; inclusion of patients with spondylolysis/listhesis
Leong et al. [13]	Anterior interbody	12.7 yrs	Adjacent level changes 52.5%; good clinical results 90%	Long-term follow-up	Surprisingly, the radiographic study showed higher % of degenerative changes than the MRI study (perhaps due to the longer follow-up)
Penta et al. [17]	Anterior interbody	10 yrs	Adjacent level changes 32%	MRI study – very sensitive	No clinical correlation

pending on the exact pathology that dominates in the individual patient. An unstable spine segment is likely to stabilise itself over a prolonged period of time, but spinal fusion is often carried out to expedite this process, provide pain relief and improve function. Thus, the non-fusion group of patients are similar to the fusion group in terms of basic pathology, and over a long-term follow-up the most important difference between the two groups will be the surgical fusion itself. A comparison is justifiable, especially if the non-fusion group has the same duration of follow-up and is age and gender matched. The present study satisfies these criteria.

The current thinking is that spinal fusion exerts its long-term effects mainly by inducing adjacent level changes in the spine, particularly at levels suprajacent to the fusion. Increased incidence of degenerative changes at levels adjacent to the fused segment have been reported by many authors [6, 8, 12, 13, 22]. The drawbacks present in these studies are summarised in Table 5. Hambly et al. [8] state that finding a control group is neither practical nor essential. This is not entirely acceptable when it is known that degenerative changes are time dependent. Comparing the results to normative data in the literature, as they suggested, is also not satisfactory. Unlike essential hypertension or diabetes mellitus, there is no threshold value to distinguish between symptomatic and asymptomatic disc degeneration. Demonstration of radiographic changes alone without clinical correlation is of little value. Even in the clinical context, the severity of symptoms is highly variable, and dependent on multitude of factors including the interpretation of 'pain' by the patients. In view of the above factors, the objective of the present study was to determine whether spinal fusion predisposes to symptomatic disc degeneration at levels above that is severe enough to cause a change in the functional outcome.

Many previous studies have used self-reported outcomes by the patients which are now known to be influenced by 'perceptual mismatch' (difference between how patients function and how they believe they function) and 'symptom sincerity' problems [5, 23]. Comparison of different series is also difficult with self-reported measures. This is where validated outcome scores come in useful. The SF36 and the Oswestry index have proven their merit in the assessment of common health problems including low back pain [4, 7, 23]. In our study there was no statistically significant difference between the fusion group and the non-fusion group in the results of the SF36 (Table 1) and Oswestry Disability Index questionnaires.

Similarly, clinical assessment of range of movement of the spine and neurological function, on their own, do not

provide sufficient information regarding functional status. Some authors have stated that functional testing is difficult due to the diverse ages and medical problems in this type of patient population [8]. In our study, performance tests could be carried out in all patients who participated in the study. The reliability and usefulness of performance tests (self-paced walk and the timed up-and-go) have been demonstrated in many studies on elderly populations, including those with medical problems [1, 9, 14, 15, 19, 20]. Once again, there was no significant difference between the fusion and the non-fusion groups in the results of self-paced walk and the timed up-and-go tests (Table 2, Table 3).

Increased incidence of radiographic changes at adjacent levels following lumbar fusion has been reported by many authors without functional correlation of these findings. Boden et al. [2] have shown that disc degeneration at an average of three levels can be seen in up to 90% of asymptomatic people over the age of 60 years. Sato and Kikuchi's [21] report shows that even the finding of radiographic instability should be correlated clinically, as the significance of 'instability' varies from patient to patient. In the present study, we found radiographic changes of instability in 14.2% of patients and disc space narrowing in 35.7% of patients who had fusion, as against a 7.4% incidence of instability and 18.5% incidence of disc space narrowing in the non-fusion group (Table 4). The incidence of radiographic changes was 50% higher in the fusion group, and this was statistically significant ($P=0.017$). The differences in radiographic changes were unaccompanied by differences in functional outcomes assessed using validated instruments.

Conclusions

Spine surgeons will be seeing increasing numbers of patients who have undergone spinal fusion for degenerative disc disease at their follow-up clinics. When faced with radiographs showing adjacent level degenerative changes, it is necessary to know their clinical impact. In this long term follow-up of 30 years, there was a significantly higher incidence of radiographic changes at adjacent levels following lumbar fusion, but this was not accompanied by a significant change in the functional outcomes assessed using validated outcome measures and performance tests.

Acknowledgements The authors gratefully acknowledge the assistance of the physiotherapy department of the Orthopaedic and Arthritic Hospital, Toronto.

References

1. Bassey EJ, Fentem PH, MacDonald IC, Scriven PM (1976) Self-paced walking as a method of exercise testing in elderly and young men. *Clin Sci Mol Med* [Suppl 51]:609–612
2. Boden S, Davis DD, Dina TS, et al (1990) Abnormal magnetic resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am* 72: 403–408
3. Chung SMK (1971) Methods of locating the 'missing patients' in long term follow-up studies. *J Bone Joint Surg Am* 53:1448–1451
4. Fairbank JCT, Couper J, Davies JB, O'Brien JP (1980) The Oswestry low back pain disability questionnaire. *Physiotherapy* 66:271–273
5. Fordyce WE, Lansky D, Calsyn DA, et al (1984) Pain measurement and pain behaviour. *Pain* 18:53–69
6. Frymoyer JW, Hanley E, Howe J, Kuhlmann D, Mattern R (1978) A comparison of radiographic findings in fusion and non fusion patients ten or more years following lumbar disc surgery. *Spine* 4:435–440
7. Garratt AM, Ruta DA, Abdallah MI, et al (1993) SF 36 health survey questionnaire. II. Responsiveness to change in health status in four common clinical conditions. *Qual Health Care* 3: 186–192
8. Hambly MF, Wiltse LL, Raghavan N, Schneiderman G, Koenig C (1998) The transition zone above a lumbosacral fusion. *Spine* 23:1785–1792
9. Harding VR, Williams AC, Richardson PH, et al (1994) The development of a battery of measures for assessing physical functioning of chronic pain patients. *Pain* 58:367–375
10. Kirkaldy-Willis WH, Farfan HF (1982) Instability of the lumbar spine. *Clin Orthop* 165:110–123
11. Lee CK (1988) Accelerated degeneration of the segment adjacent to a lumbar fusion. *Spine* 13:375–377
12. Lehmann TR, Spratt KF, Tozzi JE, Weinstein, JN et al (1987) Long term follow-up of lower lumbar fusion patients. *Spine* 12:97–104
13. Leong JCY, Chun SY, Grange WJ, Fang D (1983) Long term results of lumbar intervertebral disc prolapse. *Spine* 8:793–799
14. Okumiya K, Matsubayashi K, Nakamura T, et al (1999) The timed 'up go' test and manual button score are useful predictors of functional decline in basic and instrumental ADL in community dwelling older people. *J Am Geriatr Soc* 47:497–498
15. Peeters P, Mets T (1996) The six minute walk as an appropriate exercise test in elderly patients with chronic heart failure. *J Gerontol A Biol Sci Med Sci* 51:M147–510
16. Penny KI, Purves AM, Smith BH, et al (1999) Relationship between the chronic pain grade and measures of physical, social and psychological well-being. *Pain* 79:275–279
17. Penta M, Sandhu A, Fraser RB (1995) Magnetic resonance imaging of disc degeneration ten years after anterior lumbar interbody fusion. *Spine* 20: 743–747
18. Pope MH, Hanley EN, Mattern RE, Wilder DG, Frymoyer JW (1977) Measurement of intervertebral disc height. *Spine* 2:282–286
19. Podsaidlo D, Richardson S (1991) The timed 'up and go'. A test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc* 39:142–148
20. Purser JL, Pieper CF, Duncan PW, et al (1999) Reliability of physical performance tests in four different randomised clinical trials. *Arch Phys Med Rehabil* 80:557–561
21. Sato H, Kikuchi S (1993) The natural history of radiographic instability of the lumbar spine. *Spine* 18:2075–2079
22. Schlegel JD, Smith JA, Schlusener RL (1996) Lumbar motion segment pathology adjacent to thoracolumbar, lumbar and lumbosacral fusions. *Spine* 21: 970–981
23. Simmonds MJ, Olson SL, Jones S, et al (1988) Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. *Spine* 23:2412–2421
24. White III AA, Panjabi MM (1978) *Clinical biomechanics of the spine*, 2nd edn. JB Lippincott, Philadelphia 192–193