Zahid Askar
Douglas Wardlaw
Thillainayagamike Muthukumar
Frank Smith
Deiary Kader
Suzy Gibson

Correlation between inter-vertebral disc morphology and the results in patients undergoing Graf ligament stabilisation

Received: 7 March 2003
Revised: 6 December 2003
Accepted: 27 February 2004
Published online: 12 June 2004
© Springer-Verlag 2004

Z. Askar · D. Wardlaw (✉)
T. Muthukumar · F. Smith · D. Kader
S. Gibson
Department of Orthopaedics
and Radiology,
Grampian University Teaching Hospitals,
Woodend Hospital,
Eday Road,
AB15 6ZQ Aberdeen, Scotland
Tel.: +44-1224-556774,
Fax: +44-1224-555662,
e-mail:
Doug.Wardlaw@arh.grampian.scot.nhs.uk

Abstract *Background.* Previous studies have shown Graf ligament stabilisation procedure to give mixed results in the short to medium term. The aim of this study was to correlate the pre-operative state of the disc, multifidus muscles, age of the patient, levels operated and the clinical outcome after a mean follow-up of 47 months.

Methods. Graf ligament stabilisation procedure was carried out in 38 patients between 1996 and 1999. Their post-operative status was assessed using MacNab criteria. The post-operative follow-up was by postal questionnaires and review of the clinical notes. Disc morphology and multifidus muscle wasting was graded blindly and independently. The intra- and interobserver reliability was measured with kappa score and classified using the kappa classification of Landis and Koch. Correlation was measured with the help of Spearman correlation coefficient.

Results. Thirty-eight patients (100%) returned the questionnaires. Mean follow-up time was 47.55 months. Fifty-nine levels were operated on. Mean age was 39.68 years. The overall re-operation rate was 15.8%. The intra- and interobserver reliability was graded as good to substantial. Twenty-two patients (57.89%) were satisfied with the procedure. There was no statistically significant correlation between disc morphology, multifidus muscle wasting, sex, age, number of levels operated, the levels operated, and the satisfaction rate.

Conclusions. The indications of Graf ligament stabilisation procedure are not clear. Further work is necessary to clearly identify the indication for the procedure.

Keywords Graf ligamentoplasty · Woodend classification · Disc-state · Multifidus muscle

Introduction

The concept of ligament stabilisation was introduced by Graf [5] and advocated in patients with chronic back pain as a less invasive procedure than spinal fusion. It was widely thought to allow some motion to remain at the segments. It is recommended in patients who have mild to moderate disc degeneration after failure of the conservative forms of treatment. Grevitt et al. [6] found that the results were uniformly bad in patients with severe narrowing of the inter-vertebral disc space and spondylolisthesis

and advised against the use of this method in these patients. The rationale of this procedure is to use a posterior preconstraint to restore lumbar lordosis and stabilise the posterior facet joints in extension, thus minimizing rotatory movement and “distraction” of the facet joints. However, this mechanism is not completely understood [5].

This procedure gives an alternative to spinal fusion in patients who have only mild or moderate inter-vertebral disc space narrowing on radiographs. Kanyama et al. [10] showed that it allowed some movement at the motion segment, thus potentially avoiding problems with the adjacent levels. This is in contrast to the fusion techniques,

which completely limit motion at the segments, thus running the risk of shifting the load to the adjacent levels with the risk of future problems. In a well-selected group of patients, Kanayama et al. [10] showed that it decreased the risk of adjacent-segment deterioration compared with posterolateral fusion with instrumentation.

The first report of the procedure was by Graf himself. About 80% of his 120 patients were positively satisfied with the result after a follow-up of 6–24 months [5]. No formal independent assessment of the patients was made. Later results have been mixed. In 1998, Hadlow et al. [8] reported a significantly higher revision rate at 2 years. They also noted a poor outcome after revision to fusion.

Unfortunately, the indications for the procedure are not clearly described. We present the result of a consecutive series of our patients who underwent Graf ligament stabilisation procedure. We have correlated our clinical results with the pre-operative condition of the discs, including the state of the disc and its height, the presence of posterior annular tears, the state of the multifidus muscles (MFM), the age and sex of the patient and the number of levels operated to find out any statistically significant correlation between these factors.

Materials and methods

Graf ligament stabilisation procedure was carried out on 38 patients, between 1996 and 1999, who were suffering from dominant low back pain, some of whom had radicular-type leg pain. These patients initially underwent at least 6 months of conservative treatment including intensive physiotherapy. Patients who failed to respond to physiotherapy were treated according to the protocol set up for our spinal service, with facet injections to control their back pain. Patients continuing to have significant symptoms of low back pain after facet joint injections were considered for surgery. Prior to surgery, they had a magnetic resonance imaging (MRI) scan of the lumbar spine followed by discography to evaluate the symptomatic levels.

Patients with severe narrowing of the inter-vertebral disc space, previous history of disc enucleation or any degree of spondylolisthesis were not offered the procedure and were considered for a spinal fusion. The procedure effectively locks the facets in extension which narrows the spinal and root canals as demonstrated by Porter et al. Therefore, patients whose MRI scan showed evidence of lateral recess stenosis with or without symptoms of radicular leg pain had a decompression carried out at the affected level because of the risk of inducing symptoms by the procedure. Decompression consisted of medial facetectomy only.

All the procedures were performed by or under supervision of the senior author (DW). They were generally carried out through a midline approach. The pedicle screws were placed to avoid facet joint interference with the implants. Their positions were confirmed with peri-operative radiographs. Any decompression required was carried out through the same incision. Post-operatively, the patients were mobilised on the first or second post-operative day as and when comfortable in a lumbosacral corset for the first 3 months. They were reviewed regularly in the clinics at 3, 6, 12 and 24 monthly intervals.

Their pre-operative status was assessed from the case records. Their post-operative status was assessed by postal questionnaires based on MacNab criteria [13] (Table 1). Satisfaction with the results of the procedure was noted. Patients who had a revision dur-

Table 1 MacNab criteria (after MacNab [10])

Grade	Description
Excellent	No Pain Full activity with work
Good	Occasional pain, not interfering with work
Fair	Pain occasionally, interfering with work
Poor	Pain persists, frequently interfering with work

ing the follow-up period were classed as “poor”. Their clinical notes were studied for any peri-operative or post-operative complications. This clinical assessment was made independently by two of the co-authors (SG and DK).

The MRI scans were performed on a Siemens 0.2-T Viva Open Scanner with Numaris software (Siemens, Munich, Germany). They included sagittal (TR=3,900, TE=134) and axial T2-weighted images (TR=3,000, TE=106) from the first lumbar level through the first sacral level. The condition of the disc was classified by Woodend classification (Fig. 1) [20]. The disc height was measured by Pope’s method [17]. The state of the MFM was graded according to the Woodend classification of MFM [9]. These assessments were repeated at 4 weeks to check for the reliability of the reading. This assessment was made independently and blindly by two of the co-authors (ZA and TM).

The intra- and interobserver agreement was measured using the weighted Cohen kappa statistic package. The scores were classified into six categories based on the kappa classifications of Landis and Koch [12]. The correlation between different factors was measured using Spearman’s correlation coefficient (r) and graded according to the criteria set out by Colton [2]. A two-tailed test was also used to test the statistical significance of these findings ($\alpha=0.05$).

Results

There were 38 patients in total – 23 females (60.52%) and 15 males (39.48%). The mean age was 39.68 ± 11.53 years (range 14.02–63.49) and follow-up was 47.55 ± 9.99 months (range 24–65). The details of the levels are listed in Table 2.

Four patients had a prior decompression at the same level. Fifteen patients underwent simultaneous decompression at the time of index operation. The average kappa scores for intra- and interobserver reliability for various levels are listed in Table 3. The clinical results according to the MacNab criteria are listed in Table 4. Twenty-two patients (57.89%) were satisfied with the result of the procedure, while 16 (42.11%) were not satisfied. There was no difference in outcome comparing patients who did and did not have a primary decompression carried out. Satisfaction rate had a statistically significant correlation with the MacNab score ($r=0.85$, $P=0.001$).

Three patients were converted to fusion because of persistent symptoms of chronic low back pain. In two patients, the procedure had to be revised because of the malposition of the screws. In one patient who had back pain, the screw had penetrated through the lateral cortex, while in another the screw had perforated the medial wall of the pedicle, causing root symptoms. One patient had to undergo decompression when he developed neurological signs post-operatively. On exploration, there was no evi-

Fig. 1 The Woodend classification of MRI disc degeneration

Grade	Disc Changes			Disc Height (H)
	Axial Sections	Sagittal view	Description	
Normal (DC-1)	White (Hypertensive)		Flat or slightly convex, posterior annulus, nuclear cleft, on sagittal views. Annular margins well defined (bean shaped or rounded), no tears in axial views.	Normal (H1)
Mild (DC-2)	White or speckled		Flat or bulge of posterior annulus on sagittal views. Distortion of the bean shape, or a rounded appearance of the annulus. Small radial tears not reaching the PLL on the axial views.	Reduced by 10% (H2)
Moderate (DC-3)	Speckled or dark		Bulge or Prolapse of the posterior annulus on the sagittal views. Ill defined appearance of the annulus in axial views. Radial tears extending upto torn PLL on the sagittal/axial views, +/- prolapse or end plate changes.	Reduced by 10% - 50% (H3)
Severe (DC-4)	Dark (Hypointense)		No difference between the appearance of annulus and nucleus, +/- complex tears, +/- Prolapse with or without end plate changes.	Reduced by 50% or more (H4)

Table 2 Details of the levels

Number of levels	Number	Levels	Number
One level	20	L5/S1	25
Two levels	15	L4/L5	26
Three levels	3	L3/L4	8
Total	38	—	59

dence of screw malposition. The overall re-operation rate was 15.8%. There were no instances of superficial or deep infection in the series.

The correlation coefficients (r) and their statistical significance are listed in Table 5. None of these factors studied had a statistically significant correlation with the satisfaction rate. The condition of the disc has a “fair” degree of correlation with the satisfaction rate, however this did not reach a statistical significance.

Table 3 Average kappa scores for intra- and interobserver reliability

	Intra-observer (ZA)	Intra-observer (TM)	Inter-observer
Disc changes L5/S1	0.66	0.68	0.63
Annular tear L5/S1	0.83	0.84	0.79
Disc height L5/S1	0.67	0.67	0.56
MFM L5/S1	0.69	0.58	0.53
Disc changes L4/L5	0.68	0.68	0.62
Annular tear L4/L5	0.65	0.78	0.63
Disc height L4/L5	0.78	0.81	0.75
MFM atrophy L4/L5	0.63	0.78	0.64
Disc changes L3/L4	0.78	0.65	0.58
Annular tear L3/L4	0.72	0.64	0.57
Disc height L3/L4	0.64	0.74	0.63
MFM atrophy L3/L4	0.63	0.78	0.58

Table 4 Results

MacNab criteria	Post-operative
Excellent	2 (5.3%)
Good	15 (39.5%)
Fair	6 (15.8%)
Poor	15 (39.4 %)

Table 5 Various factors and their correlation with the satisfaction rate

Factor	Correlation coefficient (<i>r</i>)	<i>P</i> -value ($\alpha=0.05$)
Age	0.046	0.780
Sex	0.042	0.620
Length of follow-up	0.049	0.775
Number of levels	0.196	0.245
Level	0.203	0.236
Disc changes	0.303	0.068
Disc height	0.270	0.106
Annular tears	0.071	0.676
MFM atrophy	0.168	0.321

Discussion

Graf proposed the concept of Graf ligament stabilisation and recommended this procedure in patients who had "abnormal gaping of the facet joint "on twisting [5]. This was best demonstrated by the "twist test," a computed tomographic scan taken through the facet joint articulation while the patient twists the torso in relation to the legs. However Hadlow et al. [8] did not find this to be the major cause of pain in their group. Markwalder et al. [14] also made similar observations and found unpredictable results.

Guigui and Chopin [7] showed that if the patients had 2 or 3 criteria of instability (hypermobility,olisthesis of 2 mm or more on the flexion X-rays or rotatory subluxation), destabilisation occurred at the adjacent levels. They also showed that the post-operative low back pain associated with spinal stenosis did not improve if Graf ligament stabilisation was added to the decompression.

The early results of the procedure were promising. In the initial report by Graf [5], 80% of his first 120 patients were "positively satisfied" with the results of the procedure with a follow-up period between 6 and 24 months. Grevitt et al. [6] reported excellent or good results in 72% of patients, with a 34% revision rate in 50 patients after a mean follow-up of 24 months using Oswestry disability score.

Hadlow et al. [8] reported a 22% rate of good or excellent results using LBOS and a 43% rate of excellent or good subjective relief from the surgery. They concluded that outcome after Graf ligament stabilisation was associated with a worse outcome at 1 year and a significantly

higher revision rate at 2 years. They noted that revision was associated with a poor outcome similar to that seen in revision after fusion.

Rigby et al. [18] also showed that 41% of their patients were not satisfied with the procedure after a mean follow-up of 51 months. They cautioned against the continued use of this method. More recently, Gardner and Pande [4] reported a 62% rate of "excellent" or "good" subjective results in 31 patients after a mean follow-up of 7.4 years.

Our results are similar to those reported by Rigby et al. [18] and Hadlow et al. [8]. We have used Woodend MRI classification of degenerative disc disease on T2-weighted images [20] since 1999. There was substantial intra- and interobserver agreement. We graded the MFM on the basis of T2-weighted MRI using a grading system used in our hospital [9]. There was good to substantial intra- and interobserver agreement (Table 3). We excluded patients with severe disc narrowing on radiographs. This was in line with the general recommendation about the procedure.

Markwalder et al. [15] summarised the indications of the procedure as: (1) absence of arthritic changes of the facet joints, (2) an intact disc and/or only mild loss of inter-vertebral height, (3) well-trained low-back muscles, and (4) a clear-cut pain relief on test-anaesthesia of articular nerves and trial immobilization in a plastic jacket.

However we were not able to find any correlation with any of the factors stated above. In addition, symptomatic facet joint arthrosis is difficult to diagnose precisely. Different levels of facet joint arthrosis can be seen on MRI, however the clinical implication of this is unknown. The "abnormal gaping of the facet joints" as proposed by Graf is not proven, as shown by Markwalder and Merat [15].

There remain no clear-cut indications for this procedure apart from a symptomatic level producing back pain. Considering the long-term problems with spinal fusion, the future of surgical management of low back pain is perhaps in the preservation of motion rather than its restriction. Hadlow et al. [8] have been the only authors to perform a study that directly compares the outcome of Graf ligamentoplasty with posterolateral fusion in groups of "comparable" patients. Despite the various shortcomings in their study, as pointed out by Gardner et al. [3], they could not show any statistically significant difference in the outcome between the two groups. They had a higher rate of revision in the Graf ligamentoplasty group. The higher revision rate might have been because of the surgeon's readiness to revise the procedure. Moreover they did not mention how these patients fared after the procedure.

Logically, one could argue that if the results are the same, why not use a less invasive procedure? There is less muscle dissection and retraction and therefore less muscle denervation and devascularisation than posterolateral fusion. It does not invade the spinal canal risking additional nerve injury like a PLIF fusion. Finally, it avoids the po-

tential complications of anterior fusion. The Graf ligamentoplasty has been shown to decrease the risk of adjacent-segment deterioration compared with posterolateral fusion by allowing some movement across the stabilised segments [10]. This presumably happens because the bands are made from Dacron which is not pre-stretched, and they therefore potentially stretch by 15–20% after insertion.

In the light of the current evidence, such procedures still have a role to play in the surgical treatment of painful spinal segments. The indications for spinal fusion are not clear-cut. It may be that this procedure or similar proce-

dures take the place of fusion in patients who have less severe degenerative changes on radiographs and MRI.

Conclusions

In our study we were not able to find a correlation between age, sex, levels, disc changes, annular tears, disc height, MFm atrophy and the satisfaction rate in patients undergoing Graf ligament stabilisation procedure. Further work is recommended to clearly identify the indication for the procedure.

References

1. Brechbuhler D, Markwalder TM, Braun M (1998) Surgical results after soft system stabilization of the lumbar spine in degenerative disc disease – long-term results. *Acta Neurochir* 140(6):521–525
2. Colton T (1974) *Statistics in medicine*. Little, Brown and Company, Boston
3. Gardner A, Declerck GM, Hardcastle P, Markwalder TM, Moon MS, Salanova C, Shim YS (2000) The Graf ligamentoplasty procedure. Comparison with posterolateral fusion in the management of low-back pain. *Spine* 25(2):273–276
4. Gardner A, Pande KC (2002) Graf ligamentoplasty: a 7-year follow-up. *Eur Spine J* 11 [Suppl] 2:S157–163
5. Graf H (1992) Lumbar instability. Surgical treatment without fusion. *Rachis* 412:123–137
6. Grevitt MP, Gardner ADH, Spilsbury J (1995) The Graf stabilisation system: early results in 50 patients. *Eur Spine J* 4:169–175
7. Guigui P, Chopin D (1994) [Assessment of the use of the Graf ligamentoplasty in the surgical treatment of lumbar spinal stenosis. Apropos of a series of 26 patients]. [French] *Rev Chir Orthop Reparatrice Appar Mot* 80(8): 681–688
8. Hadlow SV, Fagan AB, Hillier TM, Fraser RD (1998) The Graf ligamentoplasty procedure. Comparison with posterolateral fusion in the management of low back pain. *Spine* 23(10): 1172–1179
9. Kader DF, Wardlaw D, Smith FW (2000) Correlation between the MRI changes in the lumbar multifidus muscles and leg pain. *Clin Radiol* 55(2): 145–149
10. Kanayama M, Hashimoto T, Shigenobu K, Harada M, Oha F, Ohkoshi Y, Tada H, Yamamoto K, Yamane S (2001) Adjacent-segment morbidity after Graf ligamentoplasty compared with posterolateral lumbar fusion. *J Neurosurg* 95 [Suppl] 1:5–10
11. Konno S, Kikuchi S (2000) Prospective study of surgical treatment of degenerative spondylolisthesis: comparison between decompression alone and decompression with graf system stabilization. *Spine* 25(12):1533–1537
12. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
13. MacNab I (1973) Pain and disability in degenerative disc disease. *Clin Neurosurg* 20:193–196
14. Markwalder TM, Dubach R, Braun M (1995) Soft system stabilization of the lumbar spine as an alternative surgical modality to lumbar arthrodesis in the facet syndrome. Preliminary results. *Acta Neurochir* 134(1–2):1–4
15. Markwalder TM, Mérat M (1994) The lumbar and lumbosacral facet-syn-drome-diagnostic measures, surgical treatment and results in 119 patients. *Acta Neurochir* 128:40–45
16. Papp T, Porter RW, Aspden RM, Shepperd JAN (1997) An in vitro study of the biomechanical effects of flexible stabilisation of the lumbar spine. *Spine* 22(2):151–155
17. Pope MH, Hanley EN, Mattern RE, Wilder DG, Frymoyer JW (1977) Measurement of intervertebral disc space height. *Spine* 2(4):282–286
18. Rigby MC, Selmon GP, Foy MA, Fogg AJ (2001) Graf ligament stabilisation: mid- to long-term follow-up. *Eur Spine J* 10(3):234–236
19. Skinner I, Mather-Brown N, Hardcastle P (1998) Comparative study of Graf stabilization and posterior fusion in the treatment of chronic symptomatic mechanically unstable low back syndrome. *J Musculoskeletal Res* 2(2):89–100
20. Smith F, Reggie A, Wardlaw D (2001) Painful disc lesion: can modern bi-planar magnetic resonance imaging replace discography. *British Orthopaedic Association, Annual Congress, Birmingham, 12–14 September 2001*