Review Article

Posterior inter-body fusion in lumbar spine surgery: A systematic review

Salah Fallatah, FRCSC

Department of Surgery, Umm Al-Qura University, Makkah Almukarramah, Kingdom of Saudi Arabia

Received 26 October 2012; revised 24 November 2012; accepted 7 December 2012

KEYWORDS
Degenerative lumbar spine; Posterior inter-body fusion; Posterolateral fusion

Abstract  Background: Posterolateral fusion (PLF) is a common method for achieving fusion in lumbar spine surgery. Posterior lumbar inter-body fusion (PLIF) has been reported to give a higher fusion rate and a better functional and clinical outcome.

Objectives: The objective of the review was to determine whether PLIF is more effective than PLF in improving clinical and radiological outcomes in adults surgically treated for degenerative lumbar spine conditions.

Methods: Electronic databases, bibliographies and relevant journals were searched systematically, and a meta-analysis was conducted.

Results: Of 2798 citations identified, five studies met our inclusion criteria; none was a randomized controlled trial. A total of 148 patients had undergone PLIF (intervention) and 159 PLF (control). Pooled meta-analyses showed that non-union rates were lower in the intervention group (relative risk (RR), 0.22; 95% confidence interval (CI), 0.08–0.62). The intervention group had significantly heavier disc height (weighted mean difference, 3.2 mm; 95% CI, 1.9–4.4) and less residual percentage slippage (weighted mean difference, 6.3%; 95% CI, 3.9–8.7) at final follow-up. No significant difference in segmental or total lumbar lordosis was seen. Because of the heterogeneity of the results, no conclusion could be drawn about the functional benefits of the two procedures.

Conclusion: This review suggests that PLIF results in a higher fusion rate and better correction of certain radiographic aspects of deformity than PLF. It also showed a slight but nonsignificant trend towards a better functional outcome with PLIF. The lack of randomized controlled trials and the methodological limitations of the available studies indicate the need for a sufficiently large, methodologically sound study with clinically relevant outcome measures. Until this has been done, the evidence for the beneficial effects of posterior inter-body fusion should be interpreted with caution.

© 2013 Taibah University. Production and hosting by Elsevier Ltd. All rights reserved.
Introduction

Posterolateral spinal fusion (PLF) is a long-established procedure for treating various degenerative disorders of the lumbar spine. Recently, other techniques have been described for fusion of the lumbar spine, including posterior lumbar inter-body fusion (PLIF) and unilateral transformaminal posterior lumbar inter-body fusion (TLIF). The addition of inter-body fusion allows decompression of the exiting nerve root by distracting the collapsed disc space and optimizes fusion in the load-bearing vertebral bodies with a rich blood supply. Inter-body fusion can be performed by an anterior or a posterior approach. PLIF is more technically demanding and is associated with a higher rate of complications than posterolateral fusion alone; it also adds time and cost to the procedure. A few studies have been conducted to compare PLF and PLIF in the treatment of degenerative lumbar spine conditions; however, the small sample sizes and the use of different methods to assess outcome have limited the clinical relevance of the findings.

The objective of this systematic review was to determine whether the addition of inter-body fusion to posterolateral fusion improves the clinical and radiological outcomes of adults undergoing surgical treatment for degenerative conditions of the lumbar spine, by comparing PLF with PLIF with regard to fusion, radiological correction of deformity, functional outcome and complications.

Material and Methods

Literature search

A computerized search of the electronic databases EMBASE (1980–2006) and Ovid Medline and PubMed (1966–February 2006) was performed. Hand searches of the European Spine Journal, Spine and The Journal of Spinal Disorders and Techniques were also conducted, and the reference lists of the identified studies and reviews were examined to identify further studies.

Inclusion criteria

The inclusion criteria were: (i) adult patients undergoing surgery for degenerative conditions of the lumbar spine, excluding tumour trauma and infection, with a minimum follow-up of 2 years, (ii) PLF with or without instrumentation was compared PLIF or TLIF with or without instrumentation and (iii) the outcome measure was the functional outcome of patients.

Assessment of study quality

Each published study was assessed for the quality of the study design on the Ottawa–Newcastle 8-point scale for assessing non-randomized studies. This scale grades the reporting of studies by representativeness of sample, baseline factors, assessment of outcome, statistical analysis or study design, and length of follow-up.

Data extraction

From each eligible study, data were extracted and checked the accuracy. The data consisted of the sizes and demographic data of the intervention and control groups, type of fusion, underlying diagnoses, length of follow-up, loss to follow-up, fusion rate, radiological parameters and clinical outcomes at final follow-up.

Data analysis

Because of the variety of clinical outcomes, surgical results were predefined as satisfactory if the patient had a score of <40 on the Oswestry index or >7 on the Prolo scale, a >40% gain in the Beaujon score or if the final outcome was rated as excellent or good. An outcome rating of ‘excellent’, ‘good’, ‘significantly better’, ‘satisfied’ or ‘success’ was considered satisfactory, whereas ratings of ‘fair’, ‘poor’, ‘same’, ‘worse’, ‘slightly satisfied’, ‘slightly dissatisfied’ or ‘unsuccessful’ were classified as unsatisfactory clinical outcomes.

For each study, the abstracted data were entered into the Review manager 4.2 software for statistical analysis. Pooled relative risks for dichotomous variables (complication, non-union or poor outcome) and weighted mean differences of continuous variables (final disc space height and percentage slip of spondylolysis) were calculated with a random-effects model and used to compare PLF and PLIF. The statistical heterogeneity of the pooled studies was analysed with the Higgins I2 test and was considered significant at $p < .1$.

Results

Study identification

The literature search identified 2798 potentially relevant citations, 1982 from Medline and 816 from EMBASE. Application of the eligibility criteria eliminated all but five articles for our study, four retrospective and one prospective non-randomized trial. Isthmic spondylolysis was the pre-operative diagnosis in four studies and degenerative disc disease, recurrent disc herniation, spondylolysis and spinal stenosis were the indications for surgery in the fifth. The sample size ranged from 35 to 100 patients. A total of 307 patients were evaluated, 148 of whom underwent PLIF (intervention) group and 159 PLF (control). All patients had been followed-up for a minimum of 2 years. The details of the included studies are summarized in Table 1.

Study quality

Only one study stated clearly that the cases represented all the patients who had undergone the intervention during the study period. The only prospective study did not give details of the representativeness of the sample or baseline factors, did not use a validated outcome assessment scale and did not adequately describe the surgical details or the study design and statistical analysis. Validated outcome assessment scales were used in only one study. The average follow-up period was 2–3 years in all but one study, which had a 6-year follow-up. None of the studies met the criteria for high quality on the Ottawa–Newcastle scale. The patient-specific functional outcome evaluation tools used included: the Oswestry disability index, the Prolo economic and functional scale, Beaujon score, modified Somatic Perception Questionnaire, Zung depression scale and Kirkaldy–Willis criteria.
Nonunion

Two studies defined solid fusion as the formation of crossing bony trabeculae and motion < 4 on flexion-extension on radiographs. Madan et al. used these criteria to define union in addition to the Lenke et al. criteria for bony union, and La Rosa et al. added the absence of a halo around the implant on radiograph to define solid union. Bony fusion was graded according to the Brantigan and Steffee classification in the study of Lidar.

In the pooled results, nonunion was observed in three patients (2%) in the intervention group (PLIF) and 21 patients (13%) who underwent PLF. This difference was statistically significant (p = .002, RR = 0.21, 95% CI = 0.08, 0.56; Figure 1).

Radiological correction of deformity

Four studies evaluated radiological correction of deformity (7–10), each with a different method. The intervention group had significantly higher disk height (weight mean difference = 3.2 mm, 95% CI = 1.9, 4.4) and residual percentage slippage (weighted mean difference = 6.3%, 95% CI = 3.9, 8.7) at final follow-up. No significant difference was found in segmental or total lumbar lordosis.

Functional outcomes

The instruments used to assess functional outcome are summarized in Table 1. Multiple validated outcome assessment scales were used in only one study, which was also the only one in our review that showed a significantly better functional outcome in the control group (PLF). Although the other four studies found better functional outcomes in the intervention group, none had sufficient numbers to show statistical significance. On the basis of pre-specified definitions for satisfactory and unsatisfactory results, 120 patients (81%) in the intervention group (PLIF) and 122 patients (77%) in the control group (PLF) had a satisfactory outcome (good or excellent), indicating no difference between the two groups (Figure 2).

Complications

Dehoux reported eight cases of persistent post-operative low-back pain that required hardware removal in the control group (PLF). The method used to diagnose the cause of the pain was not reported, nor whether it was improved after hardware removal. Two complications were reported in the intervention group (PLIF): in one case there was mechanical failure due to a very short fusion, and in the other the cage could not be inserted because of a very

Table 1: Characteristics of the studies included in the review.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Method</th>
<th>Participants</th>
<th>Intervention</th>
<th>Outcome measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suk et al. [10]</td>
<td>Retrospective 3.3-year follow-up</td>
<td>76 patients aged 30–60 years, with spondylolytic spondylolisthesis</td>
<td>40 PLF (1988–1991)</td>
<td>Kirkaldy–Willis criteria</td>
</tr>
<tr>
<td>Madan and Boeree [9]</td>
<td>Retrospective 2.1-year follow-up</td>
<td>44 patients (27 male, 17 female) aged 24–67 years, with isthmic spondylolisthesis</td>
<td>21 PLIF + PLF</td>
<td>Oswestry disability index; Zung depression scale; MSPQ; VAS, core set, pain drawing</td>
</tr>
<tr>
<td>Lidar et al. [8]</td>
<td>Retrospective 2-year follow-up</td>
<td>100 patients (60 male, 40 female) aged 42–50, with degenerative disc disease (48), recurrent disc herniation [22], spondylolisthesis [16], spinal stenosis [14]</td>
<td>45 PLIF + non-instrumented fusion</td>
<td>Prolo economic and functional scale; complications, fusion rate</td>
</tr>
<tr>
<td>Dehoux et al. [6]</td>
<td>Prospective non-randomized 6-year follow-up</td>
<td>52 patients aged 14–63 years with isthmic spondylolisthesis</td>
<td>27 PLIF</td>
<td>Modified Beaujon score; fusion rate</td>
</tr>
</tbody>
</table>

PLF, posterolateral fusion; PLIF, posterior inter-body fusion.

Figure 1: Fusion rates with inter-body fusion (intervention) and posterolateral fusion (control).
narrow canal. These two complications could have been avoided by careful pre-operative planning, and the descriptions of the eight cases of persistent back pain were incomplete; therefore, these 10 complications were eliminated from the pooled analysis.

In the pooled analysis (Figure 3), there was no significant difference between the two groups in the rate of complications, with nine (6%) in the intervention group (PLIF) and 10 (6.2%) in the control group (PLF).

Discussion

This review shows that PLIF improves disc height and reduces slip percentage, with a tendency to loss of correction over time. It has been reported previously that poor sagittal balance post-operatively leads to adjacent segment degeneration and poor results. The review showed no difference in segmental or total lordosis. The multiple deficiencies in the five studies identified, however, made analysis difficult. I was unable to abstract enough details to pool the radiological outcomes in all studies. Correction of radiological deformities (e.g. disc height and listhesis reduction) were reported differently by Lidar and could therefore not be pooled; in one study, the only radiological parameter reported was fusion.

Madan’s retrospective review, in which multiple functional outcome assessment scales were used and detailed post-operative evaluation was performed showed better functional outcome in the group treated with PLF than with inter-body fusion. He attributed this outcome, however, to selection bias (age, sex, extent of listhesis and disc degeneration) and retraction and scarring of the nerve roots and thecal sac. The quality of the studies included in this review obviated any conclusion on these factors.

Most of the methodological criteria for research overviews were met in this review. It had explicit inclusion and exclusion criteria, included assessment of the methodological quality of the studies, demonstrated the reproducibility of selection and assessment criteria and included a quantitative analysis. Potential selection bias was eliminated by rigorously searching many databases and reference lists, and all aspects of the selection process were conducted in duplicate. The major limitation of this review is the poor quality of the studies, which affected the quality of the cumulative data. None of the studies met the criteria of high quality on the Ottawa–Newcastle scale. Bhandari et al. stated that definitive conclusions can be reached only when high-quality randomized trials are pooled.

High fusion rates have been found with inter-body fusion. Lowe et al. reported a 90% fusion rate and 85% satisfactory clinical outcome with TLIF. Although fusion is often considered a satisfactory surgical outcome, I did not find that the functional outcome was better with inter-body fusion than with posterolateral fusion, even though the former group had a higher fusion rate. This might be due in part to the quality and design of the studies, and larger sample would be required to detect a small difference. Another possible reason is the short follow-up in most of the studies; it is possible that the results might be different with longer follow-up. Although fusion is important when evaluating the functional outcome after degenerative lumbar spine surgery, it is also crucial to recognize other factors, such as confounding comorbid conditions, pre-operative diagnosis and patient selection.

Figure 2: Functional outcomes after inter-body fusion (intervention) and after posterolateral fusion (control).

Figure 3: Rates of complications after inter-body fusion (intervention) and after posterolateral fusion (control).
Unfortunately, the studies included in this review did not do so. Recent reviews on the same topic indicated a slightly better outcome and restoration of lordosis in the group treated by inter-body fusion but no clinically significant difference between PLF and PLIF.28–30 This is why we felt further reviews are needed in an attempt to identify such differences.

TLIF was developed to address some of the complications associated with PLIF.22,23 All the inter-body interventions in the studies included in this review were performed with the PLIF technique, and the complication rate was no higher than that in the control group.

Conclusion

In conclusion, this review suggests that PLIF improves the fusion rate, correction of disc height and reduction of spondylolisthesis slip percentage. There were, however, no significant differences in functional outcome, final segmental or lordotic angle or complication rates over PLF. These conclusions are limited by the poor quality of the studies, which indicates the need for sufficiently large, methodologically sound studies to assess clinically relevant end-points. Until such studies are performed, the evidence regarding the value of adding posterior lumbar body fusion to the surgical management of degenerative lumbar spine diseases should be interpreted with caution.

References