REVIEW ARTICLE

A systematic review with meta-analysis of posterior interbody fusion versus posterolateral fusion in lumbar spondylolisthesis

Xiaoyang Liu · Yipeng Wang · Guixing Qiu ·
Xisheng Weng · Bin Yu

Abstract

Purpose To compare the clinical effectiveness of posterior lumbar interbody fusion (PLIF) and posterolateral fusion (PLF) for lumbar spondylolisthesis and to collect scientific evidence for determining which fusion method is better.

Methods After systematic search, comparative studies were selected according to eligibility criteria. Checklists by Furlan and by Cowley were used to evaluate the risk of bias of the included randomized controlled trials (RCTs) and nonrandomized controlled studies, respectively. Weighed mean differences (WMDs) and risk differences were calculated for common outcomes. The final strength of evidence was expressed as different levels recommended by the GRADE Working Group.

Results Four RCTs and five comparative observational studies were identified. Moderate-quality evidence indicated that PLIF was more effective than PLF for clinical satisfaction [odds ratios (OR) 0.49, 95 % confidence limits (95 % CI): (0.28, 0.88, $P = 0.02$)]. Moderate-quality evidence showed that no significant difference was found for the complication rate [OR 2.28, 95 % CI (0.97, 5.35), $P = 0.06$]. In secondary outcomes, moderate-quality evidence indicated that PLIF improved fusion rate [OR 0.32, 95 % CI (0.17, 0.61), $P = 0.0006$]. Low-quality evidence showed that PLIF resulted in a lower reoperation rate than PLF [OR 5.30, 95 % CI (1.47, 19.11), $P = 0.01$]. No statistical difference was found between the two groups with regard to blood loss [WMD = 76.52, 95 % CI (−310.68, 463.73), $P = 0.70$] and operating time [WMD = −1.20, 95 % CI (−40.36, 37.97), $P = 0.95$].

Conclusions Moderate-quality evidence indicates that PLIF can improve the clinical satisfaction and increase the fusion rate compared to PLF. No superiority was found between the two fusion methods in terms of complication rate, amount of blood loss, and operating time for the treatment of lumbar spondylolisthesis.

Keywords Meta-analysis · Posterior interbody fusion · Posterolateral fusion · Lumbar fusion · Spondylolisthesis

Introduction

Spondylolisthesis is defined as the forward slippage of one vertebra on another. Surgical fusion is an important method of stabilizing the spine in lumbar spondylolisthesis, which is used to reduce pain and decrease disability in patients with chronic low back pain [1]. With the development of modern surgical techniques, different fusion methods are currently available [2–8]. The surgical procedures that have been advocated include anterior interbody fusion, posterior interbody fusion, posterolateral fusion, repair of the pars interarticularis, and reduction and fusion [7–11]. Posterolateral fusion (PLF) and posterior lumbar interbody fusion (PLIF) are common choices among the various techniques available for the treatment of lumbar spondylolisthesis. Many studies have compared these approaches regarding technical demands, radiological outcomes, and clinical results; however, the scientific evidence in favor of PLIF remains weak. Inconsistent outcomes among surgeons concerning the preferred method for the two fusion methods have made it difficult to reach a consensus on the

X. Liu · Y. Wang (✉) · G. Qiu · X. Weng · B. Yu
Department of Orthopedic Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No.1 Shuaifuyuan, Wangfujing, Beijing, Dongcheng District 100730, China
e-mail: ypwang133@163.com

optimal fusion technique. Thus, the goal of this study is to evaluate the effectiveness of PLIF compared with PLF and to determine whether conclusions can be made with regard to which of the two fusion methods is better for treating patients with spondylolisthesis.

Materials and methods

Literature search

Relevant randomized controlled trials (RCTs) and comparative studies were identified from January 1960 to December 2012 by searching MEDLINE, EMBASE, and the Cochrane Central Register of Controlled Trials databases as well as by manually searching the journals of Spine, European Spine Journal, and Journal of Bone and Joint Surgery from January 1990 to December 2012. Keywords and medical subject headings related to the condition and potential treatment were identified prior to initiating the search. The search strings are shown in Table 1. Minor modifications of the search strings were required between different databases. Gray literature, including books and conference papers, were collected and these studies were included if they met inclusion criteria. No linguistic restriction was imposed on the search as recommended by the Cochrane Back Review Group editorial board [12]. Two investigators independently reviewed all subjects, abstracts, and the full text of articles that were potentially eligible based on abstract review. The eligible trials were then selected according to the inclusion criteria.

Study eligibility criteria

We systematically reviewed published studies according to the following criteria: (1) subjects who were 18 years or older and who had undergone spinal fusion for lumbar spondylolisthesis; (2) the interventions were PLIF and PLF; studies were excluded if patients accepted posterolateral fusion and interbody fusion simultaneously; (3) the study reported at least one desirable outcome; (4) all included patients were followed up for at least 1 year after surgery; and (5) studies were excluded if more than 5 % of participants had an acute spinal fracture, infection, revision [13], tumor, osteoporosis, rheumatoid arthritis, degenerative kyphosis, or degenerative lumbar scoliosis exceeding 10°.

Risk of bias assessment

The checklist by Furlan [14, 15] was used to evaluate the methodological quality of RCTs. Evaluation of the non-randomized controlled studies was performed with the checklist used by Cowley [16]. One reviewer assessed the

Table 1 Strings for electronic search

Dimension	Search strings
Study design	1. Randomized controlled trial/
	2. Controlled clinical trial/
	3. Controlled.tw.
	4. Random.tw.
	5. Randomized.tw.
	6. Randomly.tw.
	7. Trial.tw.
	8. Groups.tw.
	9. Placebo.tw.
	10. Placebo$.tw.
	11. Random$.tw.
	12. Clinical trial/
	13. Comparative study/
	14. Comparison.tw.
	15. Multicenter study/
	16. Cohort study/
	17. Prospective.tw.
	18. Retrospective.tw.
	19. Or/1–18
	20. Humans.sh.
	21. Animals.sh.
	22. 21 Not (20 and 21)
	23. 19 Not 22
Spinal disorders	24. Exp spondylolysis/
	25. Exp spondylolisthesis/
	26. Exp pars interarticularis/
	27. Exp isthmic/
	28. Exp lytic/
	29. Exp lumbar instability/
	30. Exp slipped vertebral body/
	31. Exp slipped vertebral bodies/
	32. Exp slips/
	33. Exp lumbar spine/
	34. Or/24–33
Intervention	35. Exp surgical treatment/
	36. Exp instrumentation/
	37. Exp decompression/
	38. Exp fusion/
	39. Exp arthrodesis/
	40. Exp intertransverse process arthrodesis/
	41. Exp posterolateral fusion/
	42. Exp PLF/
	43. Exp facet joint fusion/
	44. Exp facet fusion/
	45. Exp posterior lumbar interbody fusion/
	46. Exp interbody fusion/
	47. Exp PLIF/
	48. Exp cage/
	49. Or/36–43
	50. 23 and 34 and 49

Sh medical subject heading, *pt* publication type, *tw* free textword, *exp* expansion

risk of bias of the included studies. The items were scored with "yes," "no," or "unsure." A Furlan score of 6 or more out of a possible 12, or a Cowley score of 9 or more out of a possible 17, was considered to reflect "high methodological quality."

Clinical relevance

Two reviewers independently assessed the clinical relevance of the included studies according to five questions that were recommended by the Cochrane Back Review Group [14]. Each question was scored "Yes" if the clinical relevance item was met, "No" if the item was not met, and "Unclear" if data were not available to answer the question. A 30 % improvement in pain scores [17] and in functioning outcomes from baseline was considered to be clinically important.

Data collection

The data were independently extracted by two reviewers who had expertise in spinal diseases, and any disagreement that arose was discussed and resolved by consensus. The data retrieved included the following items: participant characteristics, study characteristics, specific intervention, and resultant outcomes of the comparison results. The desirable outcomes measured in this analysis were classified into primary outcomes and secondary outcomes. The clinical satisfaction was considered as the primary beneficial outcome. The complication rate was considered to be the primary harmful outcome. Secondary outcomes included fusion rate, reoperation rate, operating time, blood loss, and postoperative spinal alignment. The assessment of clinical satisfaction was based on the scores of ODI and the Prolo Economic and Functional scale and the objective
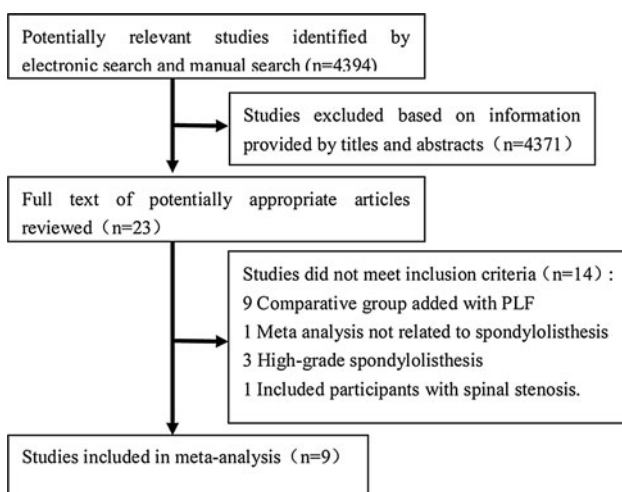

Fig. 1 Flowchart of the study selection process

evaluation from patients. "Excellent", "very good", "much better", "successful", "satisfied", or other similar appraisal was considered a satisfactory clinical result, whereas "fair", "slightly satisfied", "poor", "worse", "failure", or other similar appraisal was considered an unsatisfactory clinical result. Complications were categorized into major and minor complications according to the severity and influence on daily life. Major complications included deep infection, permanent nerve deficit, instrument failure, nonunion and revision, pulmonary embolus, and urinary tract infection with bacteremia. Minor complications consisted of superficial infection, transient nerve palsy, pain in the bone graft donor site, nonunion not revision, and urinary tract infection.

Data analysis

A meta-analysis was performed on the extracted data with RevMan 5.0 software (Cochrane IMS) using a random-effect model. For dichotomous variables, the odds ratio (OR) and 95 % confidence interval (CI) were calculated. For continuous variables, the mean difference and 95 % CI were calculated when outcome measurements in all studies were conducted on the same scale. Otherwise, the standardized mean difference and 95 % CI were calculated when the trials assessed the same outcome, but used different measurement methods. The I^2 statistics were used to test heterogeneity. An I^2 value <25 % was considered to be homogeneous, between 25 and 50 % to be of low heterogeneity, between 50 and 75 % to be of moderate heterogeneity, and above 75 % to be of high heterogeneity.

Quality of evidence

The quality of the evidence for each outcome was evaluated using a rating system with four levels recommended by the Grading of Recommendations Assessment, Development and Evaluation Working Group [14]. The level of evidence was mainly determined by RCTs. Nonrandomized controlled studies were complementary to the findings of RCTs. The initial strength of the overall body of evidence was considered "high" if the majority of RCTs were of high methodological quality. We downgraded the quality of evidence by one or two levels on the basis of the following five domains [18]: limitations of the study design [19], inconsistency [20], indirectness [21], imprecision (insufficient or imprecise data) of results [22], and publication bias [23]. The evidence was upgraded on the basis of the following criteria: (1) rated up one level when methodologically rigorous observational studies showed at least a twofold reduction or increase in risk, and rated up two levels for at least a fivefold reduction or increase in risk; (2) rated up one level for a dose–response gradient [24]. The

Table 2 Characteristics of included studies

Studies	Etiology	Participants	Intervention[a]	Outcomes
RCTs				
Cheng 2009 [28]	IS + DS	Group 1: 68 cases, mean age 48 years (38–63), follow-up rate 95.6 %, follow-up 4 years	Group 1: PLF autograft + pedicle screw/rod; decompression: laminectomy, medial facetectomy and foraminotomy; no reduction	Clinical, complication rate, fusion rate, operating time
		Group 2: 70 cases, mean age 49 years (36–62), follow-up rate 95.7 %	Group 2: PLIF autograft; (Stryker PEEK cage + pedicle screw/rod); decompression: same as PLF; no reduction	
Farrokhi 2012 [29]	IS	Group 1: 40 cases, mean age 49.66 ± 9.01 year, follow-up 1 year, total follow-up rate 100 %	Group 1: PLF decompression, instrumentation with pedicle screws (Medfield spine system II), intertransverse fusion and facet fusion by using autologous bone graft mixed with synthetic bone substitute granules (Medical Biomate Inc., Warsaw, IN)	Clinical outcomes, complication rate, fusion rate, blood loss
		Group 2: 40 cases, mean age 50.35 ± 11.30 years, follow-up 1 year	Group 2: PLIF decompression, instrumentation with pedicle screws (medfield spine system II), fusion by using autologous bone graft mixed with synthetic bone substitute granules (Medical Biomate Inc., Warsaw, IN)	
Inamdar 2006 [30]	IS + DS	Group 1: 11 cases, mean age 44.7 years (31–60), follow-up rate 90.9 %, min follow-up 11 months	Group 1: PLF autograft + pedicle screw/rod (Moss-Miami, Depuy, Warsaw); posterior decompression; no reduction	Clinical outcomes, fusion rate, complication rate, operating time, blood loss
		Group 2: 11 cases, mean age 41.4 years (21–60), follow-up rate 90.9 %	Group 2: PLIF (autograft + pedicle screw/rod); posterior decompression; no reduction	
Musluman 2011 [27]	IS	Group 1: 25 cases, mean age 47.3 years, follow-up rate 100 %	Group 1: PLF decompression, laminectomy, medial facetectomy, foraminotomy, the iliac wing as autografts; posterior transpedicular instrumentation; no reduction	Clinical, complication rate, fusion rate, revision rate and other radiological outcomes; operating time; blood loss
		Group 2: 25 cases, mean age 50.6 years (38–63), follow-up rate 100 %, total average follow-up 3.3 years (1.5–6)	Group 2: PLIF decompression, facetectomy, 2 cages (Tasarimmed) + pedicle screw/rod (Tasarimmed Spine System), spinous process autograft; with/no reduction	
Comparative studies				
Cunningham 2013 [34]	IS	Group 1: 21 cases, Mean age 46 (21–70) years, follow-up 121 (52–157) months, total follow-up rate 90 %	Group 1: PLF decompression; (autograft + TSRH pedicle screws and rods); intertransverse processes fusion	Clinical outcomes, revision rate
		Group 2: 31 cases, mean age 43 (21–72) years, follow-up 76 (45–108) months, total follow-up rate 77 %	Group 2: PLIF decomposition and reduction same as PLF; TSRH construct + two lordosing spacers (R90/Hourglass, Medtronic, Minneapolis, MN); cancellous bone graft centrally and laterally in relation to the spacers	
Dehoux 2004 [32]	IS	Group 1: 25 cases, mean age 42.4 years (14–63), minimal follow-up 6 years (75–100 months), follow-up rate 100 %	Group 1: PLF decomposition; (autograft + Cotrel-Dubousset construct); no reduction	Clinical, complication rate, fusion rate, revision rate
		Group 2: 27 cases, mean age 39.5 years (14–56)	Group 2: PLIF (Brantigan carbon-fiber cages/autograft + Steffee plates system); decompression	

Table 2 continued

Studies	Etiology	Participants	Intervention[a]	Outcomes
Dantas 2007 [31]	IS + DS	Group 1: 30 cases, mean age 52.4 years (overall age distribution 30–65), minimal follow-up 2 years, follow-up rate 100 %	Group 1: PLF Decompression: laminectomy, medial facetectomy, and foraminotomy; (autograft + pedicle screw/rod); no reduction	Clinical, complication rate, revision rate
		Group 2: 30 cases, mean age 47.6 years	Group 2: PLIF decompression: same as PLF; (titanium Plasmapore-coated intersomatic spacer/autograft + pedicle screw/rod); no reduction	
La Rosa 2001 [33]	IS	Group 1: 16 cases, mean age 57.0 years (47–67), follow-up 2–18 months, total follow-up rate 100 %	Group 1: PLF (autograft + SOCON-SRI system); decompression and reduction	Clinical, fusion rate, complication rate, reduction of slippage
		Group 2: 14 cases, mean age 57.4 years (49–67)	Group 2: PLIF (Prospace system + SOCON-SRI system); decompression; reduction: same as PLF	
Zhao 2009 [35]	IS + DS	Group 1: 14 cases, mean age 71.3 ± 4.2 years	Group 1: PLF (autograft + TSRH-3D system); decompression: laminectomy, facetectomy, and foraminotomy; no reduction;	Clinical outcome, complication rate, fusion rate, reduction of slippage
		Group 2: 16 cases, mean age 72.4 ± 5.2 years, mean follow-up 16 mo (12–24), follow-up rate 100 %	Group 2: PLIF (PEEK/autograft + TSRH-3D system); decompression and reduction: same as PLF	

Group1 PLF, *Group2* PLIF, *IS* isthmic spondylolisthesis, *DS* degenerative spondylolisthesis

[a] PLIF group accepted posterior interbody fusion without PLF

level of evidence from nonrandomized controlled studies was upgraded with caution.

Results

Search results

Flow chart shows the study selection process (Fig. 1). After the duplicate studies were excluded, a total of 4,394 studies were obtained. Based on the title and abstract, 4,371 reports were excluded because the topic of the article was not relevant to the objective of the review. In the excluded reports, two possible comparative observational studies were excluded because of a lack of complete information [25, 26]. Finally, we identified a total of nine eligible studies, consisting of four RCTs [27–30] and five comparative observational studies [31–35] that involved a total of 520 patients. The individual sample sizes ranged from 22 to 138 patients. In one comparative study, one patient younger than 18 years of age was enrolled in each group [32]. One patient with traumatic spondylolisthesis was enrolled in the PLF group from one RCT [30]. Both trials were selected because most participants were eligible despite the small effect of these ineligible cases.

Five of the studies recruited patients who had been diagnosed with isthmic spondylolisthesis [27, 29, 32–34].

The other four trials selected a mix of patients with degenerative and isthmic spondylolisthesis [28, 30, 31, 35]. Patients with a diagnosis of degenerative spondylolisthesis accounted for 27.3 % of the total patients. All participants had a history of back pain with or without radicular pain. Internal fixation was used in both fusion procedures. Detailed information on the study designs, characteristics of participants, follow-up, interventions, instruments, and outcomes are shown in Table 2.

Risk of bias of included studies

The Furlan scores for the four RCTs ranged from 5 to 9 out of 12 (Table 3). Three RCTs received Furlan scores of six or higher, indicating the overall lower risk of bias of the trials. The most notable methodological shortcomings were uncertainties regarding blinding procedures, with only one "Yes" item for the 4 RCTs. In two studies, there was a clear attempt at concealment of the group allocation and method of randomization [27, 29]. The average follow-up of the trials ranged from 1 to 4 years. The Cowley scores for the five nonrandomized controlled studies ranged from 11 to 13 out of 17 (Table 4). All nonrandomized controlled studies received Cowley scores >9 and represented "high methodological quality".

Table 3 Risk of bias assessment of included RCTs using the checklist by Furlan

Criteria	Cheng 2009 [28]	Farrokhi 2012 [29]	Inamdar 2006 [30]	Musluman 2011 [27]
A 1. Was the method of randomization adequate?	Unclear	Yes	Unclear	Yes
B 2. Was the treatment allocation concealed?	Unclear	Yes	Unclear	Yes
C Was knowledge of the allocated interventions adequately prevented during the study?				
3. Was the patient blinded to the intervention?	Unclear	Unclear	Unclear	No
4. Was the care provider blinded to the intervention?	Unclear	Unclear	Unclear	No
5. Was the outcome assessor blinded to the intervention?	Unclear	Yes	Unclear	Unclear
D Were incomplete outcome data adequately addressed?				
6. Was the drop-out rate described and acceptable?	Unclear	Yes	Yes	Yes
7. Were all randomized participants analyzed in the group to which they were allocated?	Yes	Yes	Yes	Yes
E 8. Are reports of the study free of suggestion of selective outcome reporting?	No	Yes	Yes	Yes
F Other sources of potential bias:				
9. Were the groups similar at baseline regarding the most important prognostic indicators?	Yes	Unclear	Yes	Yes
10. Were co-interventions avoided or similar?	Yes	Yes	No	Unclear
11. Was the compliance acceptable in all groups?	Yes	Yes	Yes	Yes
12. Was the timing of the outcome assessment similar in all groups?	Yes	Yes	Yes	Yes
Scores	5	9	6	8

Table 4 Risk of bias assessment of included observational studies using the checklist by Cowley

	Cunningham 2013 [34]	Dantas 2007 [31]	Dehoux 2004 [32]	La Rosa 2001 [33]	Zhao 2009 [35]
1. Method of selection of patients identified and appropriateness	Yes	Yes	Yes	Yes	Yes
2. Number of patients deceased or lost to follow-up reported or included in appropriate statistical analysis	Yes	Yes	Yes	Yes	Yes
3. Follow-up period range and mean given (minimum = N)	Yes	Yes	No	No	Yes
4. Prosthesis models specified	Yes	Yes	Yes	Yes	Yes
5. Clearly defined criteria for measuring outcomes	Yes	Yes	Yes	Yes	Yes
6. Valid statistical analysis undertaken	Yes	Yes	Yes	Yes	No
7. Data given for deceased patients (information)	No	Yes	Yes	Yes	Yes
8. Age range and mean age reported	Yes	Yes	Yes	Yes	Yes
9. Numbers of men and women given	Yes	Yes	Yes	Yes	No
10. Weight range and mean weight given	No	No	No	No	No
11. Preop diagnoses with percentages of patients given	Yes	Yes	Yes	Yes	Yes
12. Clinical evaluation independent of operating surgeon	Yes	No	No	Unclear	No
13. Radiological evaluation independent and blinded to clinical results	No	No	Yes	Unclear	No
14. Results given for specific models	No	No	Yes	Yes	Yes
15. Quantification of outcomes	Yes	Yes	Yes	Yes	Yes
16. Follow-up data compared with preoperative data (mean and range)	Yes	Yes	No	Yes	Yes
17. Independence of investigators (no vested interest) stated	Yes	No	No	No	No
Scores	13	12	12	12	11

A positive answer (yes) to any question counts as 1 point

Clinical relevance

The clinical relevance of the included studies is presented in Table 5. All of the included studies described the interventions and treatment settings with sufficient detail for clinicians to replicate the treatment in clinical practice.

Two studies did not report the important clinical end points [29, 35], such as a reduction in pain or improvement in function. In four included trials (one RCT [27] and three controlled clinical studies [31, 34, 36]), the reviewers considered the likely treatment benefits to be worth the potential harm. In five studies [27, 28, 31, 34, 36], the size

Table 5 Clinical relevance

Question	RCTs				Observational studies				
	Cheng 2009 [28]	Farrokhi 2012 [29]	Inamdar 2006 [30]	Musluman 2011 [27]	Cunningham 2013 [34]	Dehoux 2004 [32]	Dantas 2007 [31]	La Rosa 2001 [33]	Zhao 2009 [35]
1. Are the patients described in detail so that you can decide whether they are comparable to those that you see in your practice?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2. Are the interventions and treatment settings described well enough so that you can provide the same for your patients?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
3. Were all clinically relevant outcomes measured and reported?	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No
4. Is the size of the effect clinically important?	Yes	Unclear	Unclear	Yes	Yes	Unclear	Yes	Yes	Unclear
5. Are the likely treatment benefits worth the potential harms?	No	No	No	Yes	Yes	Unclear	Yes	Yes	No

of the effect was considered to be clinically important. More than 30 % improvement in back pain scores were observed in six clinical results. Information about function outcome was not provided in the study by Dehoux [32].
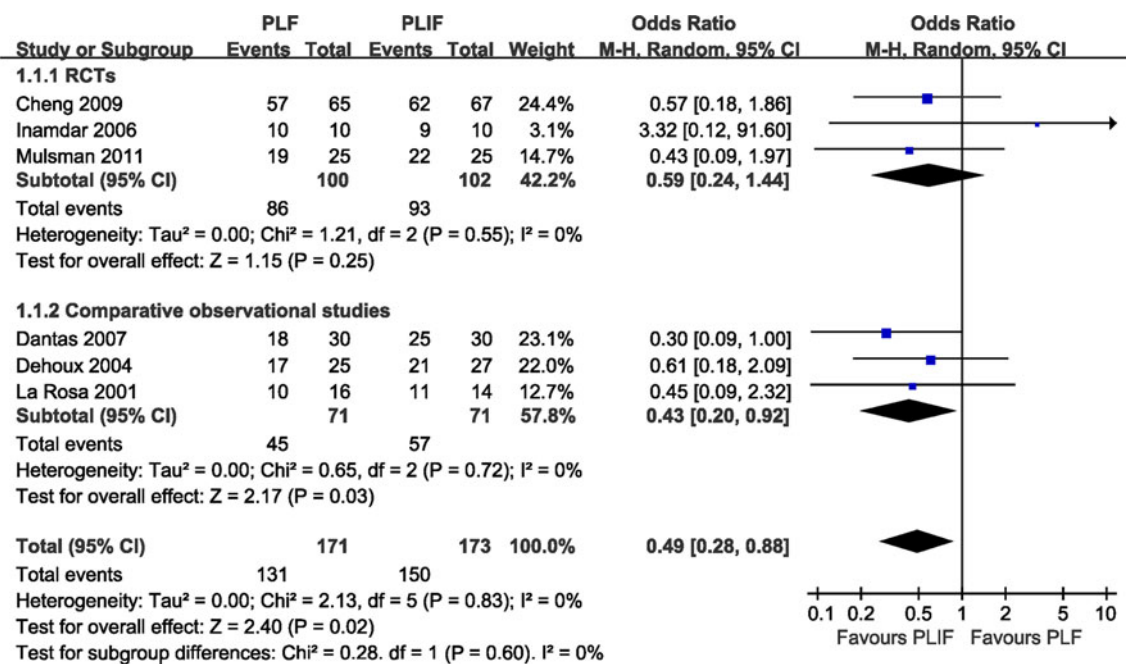
Meta-analysis results

No significant difference in demographics, symptoms, level and grade of slip, and preoperative distribution of lifestyle factors was found between the two surgical groups from the included trials. Because of imprecision of results, we downgraded one level of evidence for clinical satisfaction. There was moderate evidence from three RCTs (202 patients) and three observational studies (142 patients) that the PLIF treatment procedure was more effective than the PLF treatment for clinical satisfaction, with an OR of 0.49(95 % CI 0.28–0.88, $P = 0.02$). A total of 150 (86.7 %) of 173 patients were satisfied with the surgical outcome in the PLIF group compared to 131 (76.6 %) of 171 patients in the PLF group (Fig. 2). Subgroup analysis revealed inconsistent trends of RCTs and comparative observational studies, with the former showing a comparable outcome (OR 0.59, 95 % CI 0.24–1.44, $P = 0.25$) and the latter a significantly higher satisfaction rate with PLIF (OR 0.43, 95 % CI 0.20–0.92, $P = 0.03$). Though no significant difference was found in the subgroup of RCTs, a higher satisfaction rate was found in patients receiving PLIF.

The indexes applied to assess the improvement in symptoms and function varied among the selected trials and included the 36-Item Short Form Health Survey, Short Form (SF)-12 v2, and SF-6D R2. Statistical analysis was feasible after standardization pooling for comparing
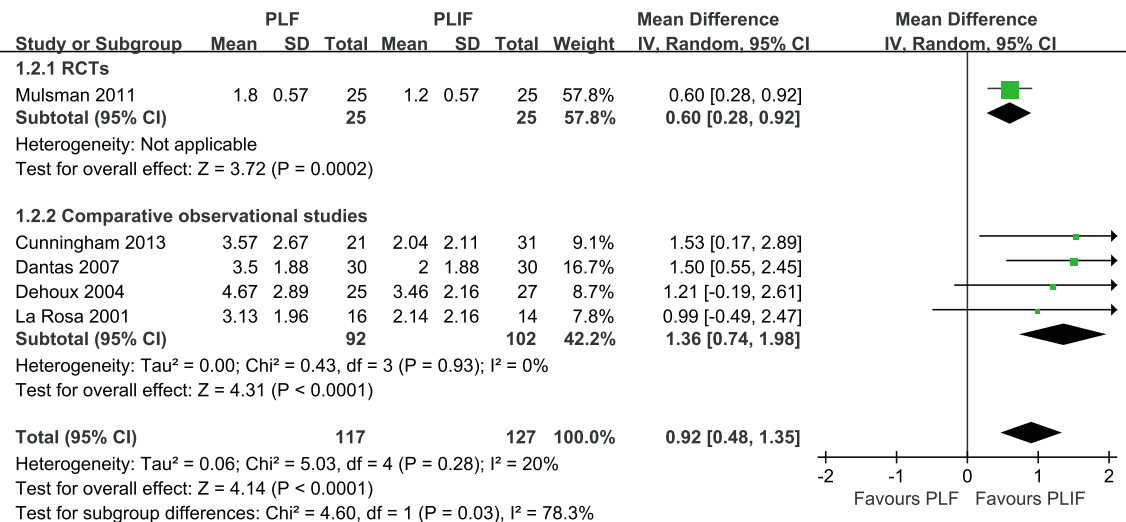
functional improvement. Improvement in functional status postoperatively was identified for both interventions, but was more significant in the PLIF group. However, the superiority of PLIF in the reduction of postoperative pain and the improvement of function decreased as the follow-up went on [27, 34]. The level of evidence was downgraded because only one RCT was included in the evaluation of postoperative back pain. There was moderate-quality evidence from one RCT (50 patients) and four observational studies (194 patients) that the PLIF was more effective than the PLF for postoperative back pain with WMD 0.92 (95 % CI 0.48–1.35, $P < 0.0001$; Fig. 3). Postoperative functional performance was assessed using the Oswestry Disability Index (ODI) questionnaire. The total score ranged from 0 to 100, in which 100 indicates the most severe disability. There was low-quality evidence from two RCTs (202 patients) and three observational studies (142 patients) that the PLIF was more effective than the PLF for improving ODI with WMD 1.30 (95 % CI 0.25–2.35, $P = 0.01$).

There was moderate-quality evidence from four RCTs (282 patients) and four observational studies (172 patients) that there was no significant difference in the complication rate [OR 2.28, 95 % CI (0.97, 5.35), $P = 0.06$], which reflects the primary harm outcome (Fig. 4). The strength of evidence was decreased due to inconsistency between RCTs and observational studies. Sensitivity analysis revealed inconsistent trends for both subgroups of RCTs and comparative observational studies, with the former showing a comparable outcome (OR 1.25, 95 % CI 0.33–4.73, $P = 0.74$) and the latter a significantly lower complication rate with PLIF (OR 4.62, 95 % CI

Study or Subgroup	PLF Events	Total	PLIF Events	Total	Weight	Odds Ratio M-H, Random, 95% CI	Odds Ratio M-H, Random, 95% CI
1.1.1 RCTs							
Cheng 2009	57	65	62	67	24.4%	0.57 [0.18, 1.86]	
Inamdar 2006	10	10	9	10	3.1%	3.32 [0.12, 91.60]	
Mulsman 2011	19	25	22	25	14.7%	0.43 [0.09, 1.97]	
Subtotal (95% CI)		100		102	42.2%	0.59 [0.24, 1.44]	
Total events	86		93				

Heterogeneity: Tau² = 0.00; Chi² = 1.21, df = 2 (P = 0.55); I² = 0%
Test for overall effect: Z = 1.15 (P = 0.25)

Study or Subgroup	PLF Events	Total	PLIF Events	Total	Weight	Odds Ratio M-H, Random, 95% CI	
1.1.2 Comparative observational studies							
Dantas 2007	18	30	25	30	23.1%	0.30 [0.09, 1.00]	
Dehoux 2004	17	25	21	27	22.0%	0.61 [0.18, 2.09]	
La Rosa 2001	10	16	11	14	12.7%	0.45 [0.09, 2.32]	
Subtotal (95% CI)		71		71	57.8%	0.43 [0.20, 0.92]	
Total events	45		57				

Heterogeneity: Tau² = 0.00; Chi² = 0.65, df = 2 (P = 0.72); I² = 0%
Test for overall effect: Z = 2.17 (P = 0.03)

Total (95% CI)		171		173	100.0%	0.49 [0.28, 0.88]	
Total events	131		150				

Heterogeneity: Tau² = 0.00; Chi² = 2.13, df = 5 (P = 0.83); I² = 0%
Test for overall effect: Z = 2.40 (P = 0.02)
Test for subgroup differences: Chi² = 0.28. df = 1 (P = 0.60). I² = 0%

0.1 0.2 0.5 1 2 5 10
Favours PLIF Favours PLF

Fig. 2 Clinical satisfaction of PLF versus PLIF for the treatment of lumbar spondylolisthesis. The assessment of clinical satisfaction was based on the scores of ODI and the Prolo Economic and Functional scale and the objective evaluation from patients. Significant difference was observed for overall effect, favoring PLIF with higher clinical satisfaction. Subgroup analysis showed inconsistent results between RCTs and observational studies

Study or Subgroup	PLF Mean	SD	Total	PLIF Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI
1.2.1 RCTs									
Mulsman 2011	1.8	0.57	25	1.2	0.57	25	57.8%	0.60 [0.28, 0.92]	
Subtotal (95% CI)			25			25	57.8%	0.60 [0.28, 0.92]	

Heterogeneity: Not applicable
Test for overall effect: Z = 3.72 (P = 0.0002)

Study or Subgroup	PLF Mean	SD	Total	PLIF Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	
1.2.2 Comparative observational studies									
Cunningham 2013	3.57	2.67	21	2.04	2.11	31	9.1%	1.53 [0.17, 2.89]	
Dantas 2007	3.5	1.88	30	2	1.88	30	16.7%	1.50 [0.55, 2.45]	
Dehoux 2004	4.67	2.89	25	3.46	2.16	27	8.7%	1.21 [-0.19, 2.61]	
La Rosa 2001	3.13	1.96	16	2.14	2.16	14	7.8%	0.99 [-0.49, 2.47]	
Subtotal (95% CI)			92			102	42.2%	1.36 [0.74, 1.98]	

Heterogeneity: Tau² = 0.00; Chi² = 0.43, df = 3 (P = 0.93); I² = 0%
Test for overall effect: Z = 4.31 (P < 0.0001)

Total (95% CI)			117			127	100.0%	0.92 [0.48, 1.35]	

Heterogeneity: Tau² = 0.06; Chi² = 5.03, df = 4 (P = 0.28); I² = 20%
Test for overall effect: Z = 4.14 (P < 0.0001)
Test for subgroup differences: Chi² = 4.60, df = 1 (P = 0.03), I² = 78.3%

-2 -1 0 1 2
Favours PLF Favours PLIF

Fig. 3 Postoperative back pain of PLF versus PLIF for the treatment of lumbar spondylolisthesis. Both overall and subgroup analyses showed statistical differences between the two procedures. Relief of back pain was more significant in the PLIF group compared to the PLF group

1.90–11.21, P = 0.0007). The analysis of heterogeneity revealed an overall I² score of 58 %, with I² scores of 71 and 0 % for the two subgroups, indicating substantial heterogeneity in the RCTs. The study by Inamdar et al. demonstrated a substantially higher complication rate with PLIF, especially postoperatively persistent back pain in four patients. A sensitivity analysis with removal of the study by Inamdar et al. [30] revealed a significantly lower complication rate in the PLIF group (OR 2.50, 95 % CI 1.27–4.95, P = 0.008). With respect to major or minor complications, there was moderate-quality evidence that there was no significant difference between two procedures [OR: 2.65, 95 % CI (0.84, 8.36), P = 0.10; OR 1.62, 95 % CI (0.59, 4.43), P = 0.35].

In the secondary outcomes, the level of evidence for fusion rate was downgraded because of the imprecision.
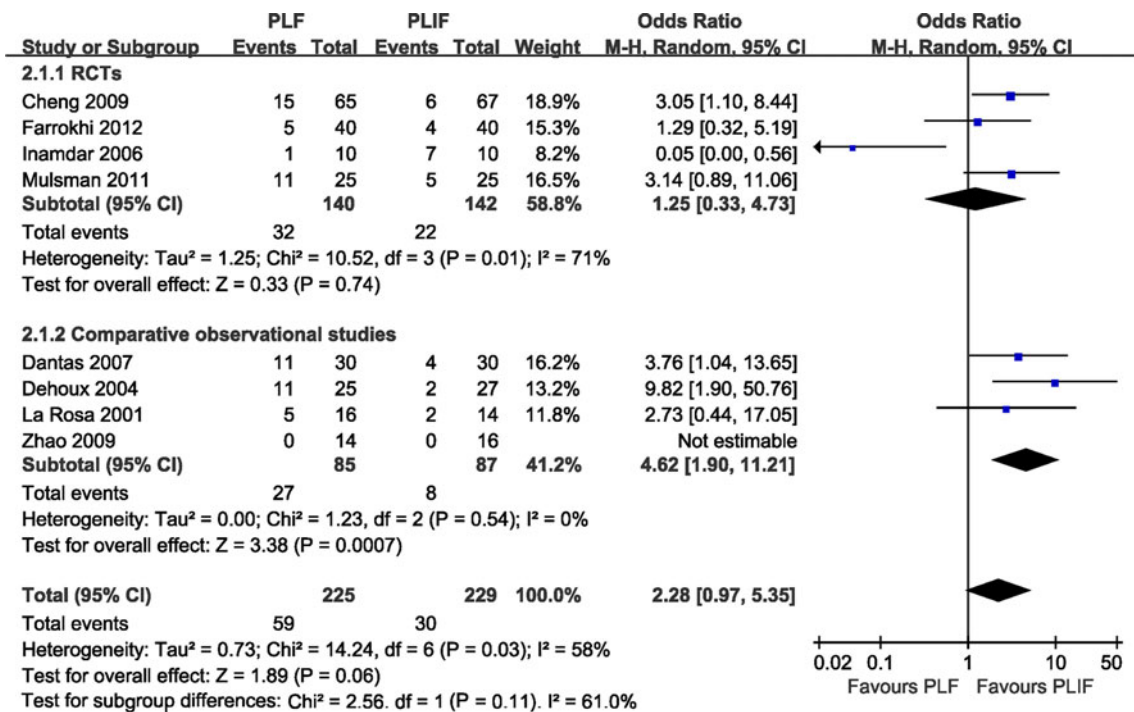
Fig. 4 Complication rate of PLF versus PLIF for the treatment of lumbar spondylolisthesis. No significant difference was found for the complication rate between the two fusion procedures. I^2 was >50 %, indicating substantial heterogeneity in the RCTs
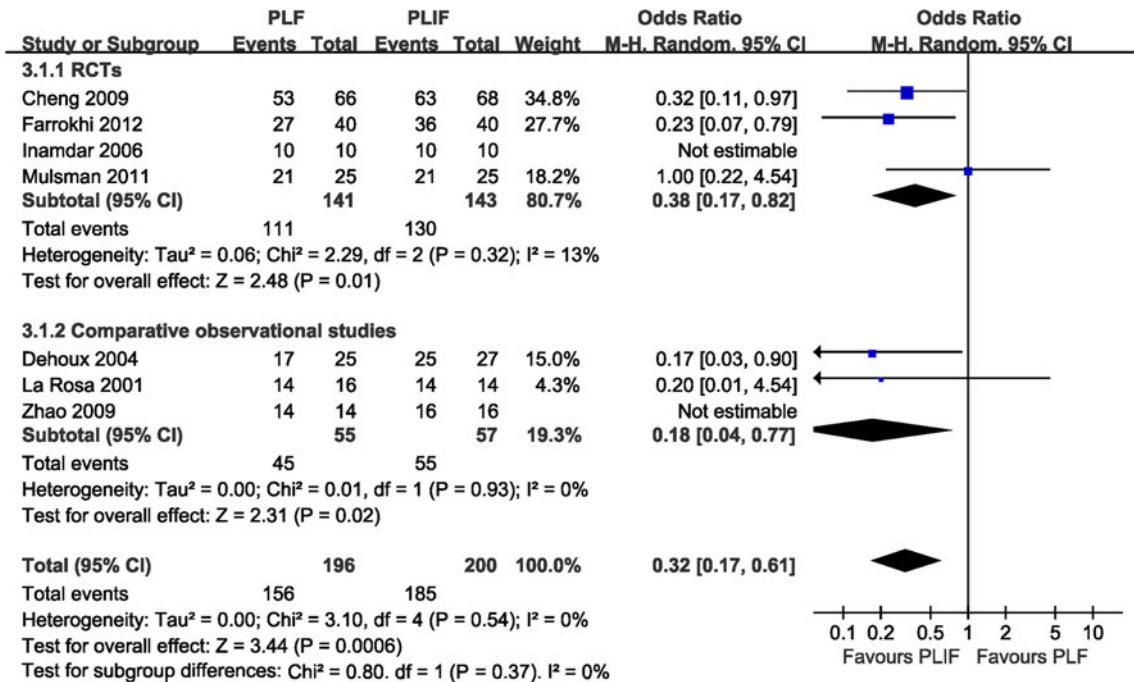
Fig. 5 Fusion rate of PLF versus PLIF for the treatment of lumbar spondylolisthesis. The fusion rate in the PLIF group was lower than that in the PLF group

There was moderate-quality evidence from four RCTs (284 patients) and three observational studies (112 patients) that PLIF was more effective than PLF for improving fusion rate, with an OR of 0.32 (95 % CI 0.17–0.61, P = 0.0006).

The fusion rates of the PLIF group and the PLF group were 92.5 and 79.6 %, respectively (Fig. 5). The strength of evidence for reoperation rate was downgraded by one level because of inconsistency between subgroups. There was

low-quality evidence from one RCT (50 patients) and four observational studies (172 patients) that the PLIF was more effective than the PLF for the reduction of reoperation rate with OR 5.30 (95 % CI 1.47–19.11, $P = 0.01$). The reoperation rates of the PLIF group and the PLF group were 3.6 and 17.3 %, respectively (Fig. 6). Subgroup analysis showed inconsistent trends of RCTs and comparative observational studies. Although no significant difference was found in the subgroup of RCTs, patients who accepted PLIF showed a lower reoperation rate.

There was low-quality evidence from 2 RCTs that there was no statistically significant difference between the two procedures with regard to blood loss [WMD = 76.52, 95 % CI (−310.68, 463.733), $P = 0.70$] and operating time [WMD = −1.20, 95 % CI (−40.36, 37.97), $P = 0.95$] (Figs. 7, 8). The strength of evidence was downgraded by two levels because of imprecision and inconsistency. The I^2 scores of 92 % indicated considerable heterogeneity in the included RCTs. The greater blood loss in the PLF group in the study by Musluman et al. [27] may result from the procedure used to collect bone from the iliac wing. For the operating time, an I^2 score of 81 % indicated substantial heterogeneity in both RCTs. The study by Cheng et al. [28] that used low-quality methodology reported a longer operating time for the PLF procedure, which could be due to the additional procedure of bone collection from the iliac wing.

A prospective randomized clinical study with a 2-year follow-up period showed that lumbar lordosis and the segmental angle were restored and maintained in the PLIF group compared to preoperative alignment ($P < 0.05$) [27]. There was low-quality evidence that PLIF significantly restored the segmental and lumbar lordotic angles, but PLF did not [27].
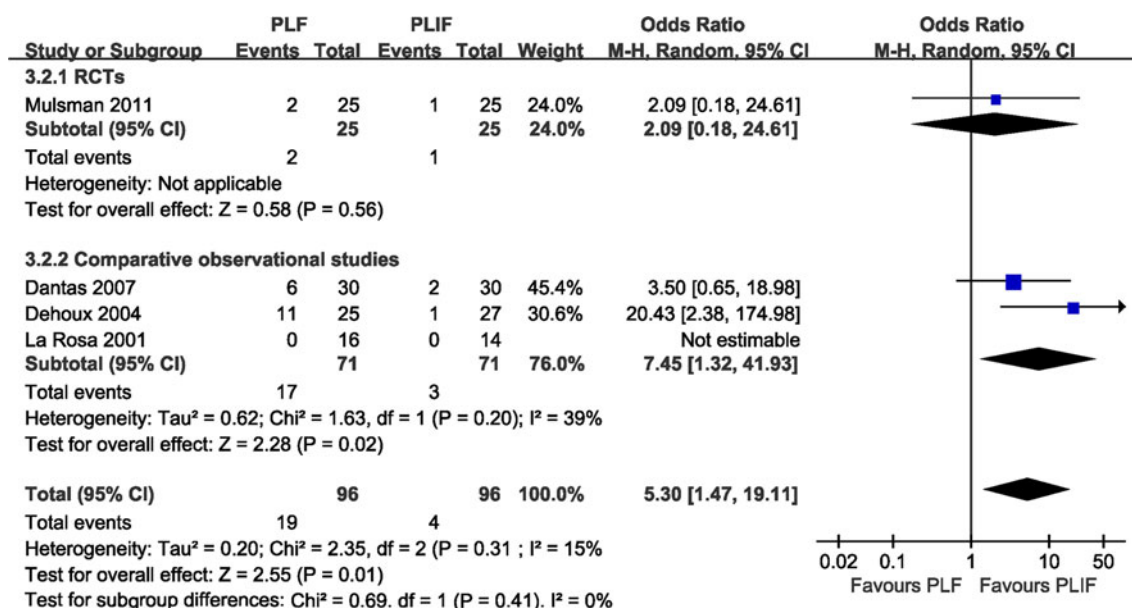
Fig. 6 Reoperation rate of PLF versus PLIF for the treatment of lumbar spondylolisthesis. Lower reoperation rate in the PLIF group was observed in this forest plot
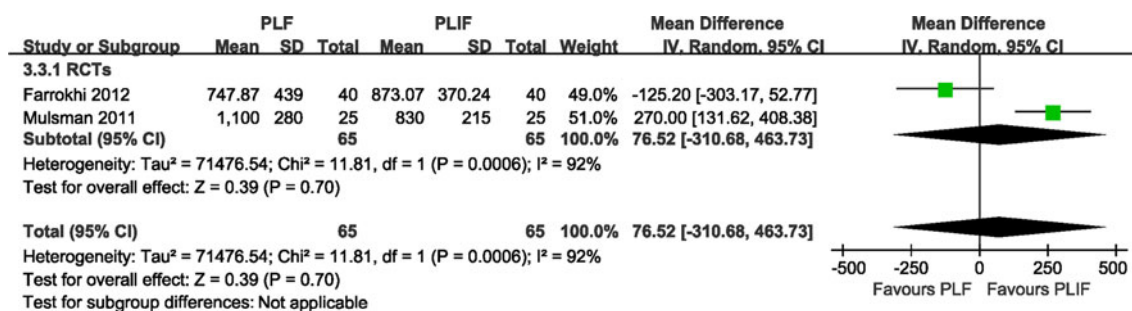
Fig. 7 Blood loss of PLF versus PLIF for the treatment of lumbar spondylolisthesis. No significant difference was found between the two groups. The I^2 was >75 %, indicating considerable heterogeneity in the included studies

Study or Subgroup	PLF Mean	SD	Total	PLIF Mean	SD	Total	Weight	Mean Difference IV. Random, 95% CI
3.4.1 RCTs								
Cheng 2009	210	62	70	192	67	68	52.0%	18.00 [-3.55, 39.55]
Mulsman 2011	146	48.75	25	168	47.5	25	48.0%	-22.00 [-48.68, 4.68]
Subtotal (95% CI)			95			93	100.0%	-1.20 [-40.36, 37.97]

Heterogeneity: Tau² = 646.88; Chi² = 5.22, df = 1 (P = 0.02); I² = 81%
Test for overall effect: Z = 0.06 (P = 0.95)

| Total (95% CI) | | | 95 | | | 93 | 100.0% | -1.20 [-40.36, 37.97] |

Heterogeneity: Tau² = 646.88; Chi² = 5.22, df = 1 (P = 0.02); I² = 81%
Test for overall effect: Z = 0.06 (P = 0.95)
Test for subgroup differences: Not applicable

Fig. 8 Operating time of PLF versus PLIF for the treatment of lumbar spondylolisthesis. No significant difference was found for operating time between the two groups. Inconsistency was obvious in the included studies

Discussion

This meta-analysis identified four RCTs and five observational studies that compared PLIF with PLF for lumbar spondylolisthesis. The findings revealed that there was moderate-quality evidence indicating that PLIF had advantages of clinical satisfaction, reduction in postoperative pain, and improvements in fusion rate compared to PLF. Low-quality evidence indicated that there was no significant difference between the two procedures for complication rate, blood loss, and operating time. Patients with PLIF reached better clinical outcomes and satisfaction than patients with PLF.

In the past several decades, the continuous modification and refinement of surgical techniques, such as minimization of the level of neural retraction required and avoidance of broad dissection of the paraspinal musculature during PLIF, have contributed to a reduction in the operative risks, operating time, and blood loss during PLIF [37]. During a PLF procedure, the broad dissection used that exceeds the facet joint may lead to a transient increase in postoperative pain, which can further influence patients' satisfaction of the procedure. PLIF can overcome these drawbacks and provide anterior column support, which helps to restore lumbar lordosis and intervertebral space height as well as increase fusion rate. A higher fusion rate with PLIF was identified in the current analysis compared to PLF. Some researchers believe that once the unstable segment is successfully fused, mechanical back pain due to a pars defect or facet arthropathy can be reduced, which may contribute to good functional outcomes [38, 39]. Therefore, successful arthrodesis will most likely indicate a satisfactory clinical outcome. However, nonunion and its associated complications may result in postoperative recurrent back pain or even failure of the surgery and reoperation, thus preventing a satisfactory outcome [40–42]. Compared to PLF, PLIF resulted in a higher fusion rate. Many studies have postulated that successful fusion status can result in better functional outcome and better satisfaction [43, 44]. This

may indicate that PLIF has a clinical advantage over PLF. However, the results from various studies are conflicting [45–48], and a complete correlation between good outcome and fusion rate was not recorded in some studies [33, 49]. Moreover, some studies have indicated that there was no significant difference in the reduction of postoperative pain between the two interventions [50].

The present meta-analysis captured information on the characteristics of included patients, detailed information of fusion procedures, and clinically important end points, which will enable clinicians in the field to determine whether the results apply to their patient population and how to apply these procedures in their clinical practice. For some clinical outcomes, statistical significance does not necessarily mean that the change is clinically important [51]. The minimal important change (MIC) can provide more information and help users evaluate the effect size of interventions. There was wide variation in criteria and statistical techniques used to define MIC in the literature and reviews. In 2008, Ostelo et al. [52] proposed MIC values of 15 for the visual analog scale and 10 for the Oswestry Disability Index. When the baseline score was taken into account, a 30 % improvement was considered to be a useful threshold for identifying clinically meaningful improvements for each of these measures. Thus, we used similar criteria to determine important clinical differences in pain reduction and functional improvement [14]. The effect size in the five studies was considered to be clinically important [27, 28, 31, 34, 36].

The combination of RCTs and comparative observational studies is becoming increasingly common for the evaluation of surgical treatments. Well-designed observational studies are believed to be beneficial complements to the findings of RCTs [53, 54], as they may dilute the selection bias of RCTs produced by the rigorous criteria for selecting participants. In our included observational studies, patients were allocated to a designated treatment group according to specific time sequence, such as the date of treatment. In addition, all of the studies attempted to

balance the intervention groups for possibly important prognostic indicators. In general, a meta-analysis based on observational studies results in low-quality evidence [55]. However, in the present analysis, we mainly obtained moderate-quality evidence after comprehensively evaluating the level of evidence.

We also found that the outcomes from RCTs may differ from those in observational studies, which could be a result of different study designs and bias. One of the RCTs was rated as having a "high risk of bias" because it met <6 of the 12 CBRG criteria and had serious flaws. However, a sensitivity analysis with removal of the low-quality study by Cheng et al. [28] showed a similar result, indicating that the heterogeneity was not from the low quality of methodology. A subanalysis found that surgical procedure may be the source of the risk of bias and heterogeneity for special outcomes [27–29]; however, it is possible that the intrinsic nature of different fusions is the source of the discrepancy. Therefore, the results of this study should be interpreted with caution, especially for reoperation rate, blood loss, and operative time.

This study had several limitations. According to our search results and inclusion criteria, four RCTs and five observational studies were included in this meta-analysis. In addition, the number of studies for each of the outcomes varied from one plot to another. The small number of included trials and incomplete data decreased the quality of evidence and the power of the subgroup analyses. There was also potential publication bias, because we only retrieved published literature in peer-reviewed journals. However, we were not able to evaluate this because of the limitation of the included studies. Blinding was also not completely performed in the four RCTs, though randomization and allocation concealment were performed in two trials [14, 27]. Inadequate blinding has been reported to produce a 15 % overestimation of treatment effects [56]. In addition, the indications for surgery and the actual type of surgical procedure varied in the two fusion methods, which may be more obvious in observational studies. Inconsistency between studies increased the risk of selection bias. In this analysis, eights included studies contained <50 subjects in the smallest group [27, 29–32, 35, 36, 57]. Studies with a small sample size can increase heterogeneity and bias. Therefore, the pooled ORs should be treated with caution. A greater number of well-designed RCTs will help provide much stronger evidence for clinical decision-making in the future.

Conclusion

PLIF had advantages of a reduction of postoperative pain as well as improvement of patient satisfaction and function compared to PLF. In addition, a PLIF can increase the fusion rate and decrease the reoperation rate. We identified low-quality evidence showing no significant difference between the two fusion procedures with regard to blood loss and operating time.

References

1. Ha KY, Na KH, Shin JH, Kim KW (2008) Comparison of posterolateral fusion with and without additional posterior lumbar interbody fusion for degenerative lumbar spondylolisthesis. J Spinal Disord Tech 21:229–234. doi:10.1097/BSD.0b013e3180eaa202

2. Kim NH, Lee JW (1999) Anterior interbody fusion versus posterolateral fusion with transpedicular fixation for isthmic spondylolisthesis in adults. A comparison of clinical results. Spine (Phila Pa 1976) 24:812–816, 817

3. Lin PM (1985) Posterior lumbar interbody fusion technique: complications and pitfalls. Clin Orthop Relat Res 193:90–102

4. DeWald RL, Faut MM, Taddonio RF, Neuwirth MG (1981) Severe lumbosacral spondylolisthesis in adolescents and children. Reduction and staged circumferential fusion. J Bone Joint Surg Am 63:619–626

5. Deguchi M, Rapoff AJ, Zdeblick TA (1998) Posterolateral fusion for isthmic spondylolisthesis in adults: analysis of fusion rate and clinical results. J Spinal Disord 11:459–464

6. Rombold C (1966) Treatment of spondylolisthesis by posterolateral fusion, resection of the pars interarticularis, and prompt mobilization of the patient. An end-result study of seventy-three patients. J Bone Joint Surg Am 48:1282–1300

7. Gjessing MH (1951) Osteoplastic anterior fusion of the lower lumbar spine in spondylolisthesis, localized spondylosis, and tuberculous spondylitis. Acta Orthop Scand 20:200–213

8. Cloward RB (1981) Spondylolisthesis: treatment by laminectomy and posterior interbody fusion. Clin Orthop Relat Res 154:74–82

9. Schnee CL, Freese A, Ansell LV (1997) Outcome analysis for adults with spondylolisthesis treated with posterolateral fusion and transpedicular screw fixation. J Neurosurg 86:56–63. doi:10.3171/jns.1997.86.1.0056

10. Noggle JC, Sciubba DM, Samdani AF, Anderson DG, Betz RR, Asghar J (2008) Minimally invasive direct repair of lumbar spondylolysis with a pedicle screw and hook construct. Neurosurg Focus 25:E15. doi:10.3171/FOC/2008/25/8/E15

11. Tsutsumimoto T, Shimogata M, Ohta H, Misawa H (2009) Mini-open versus conventional open posterior lumbar interbody fusion for the treatment of lumbar degenerative spondylolisthesis: comparison of paraspinal muscle damage and slip reduction. Spine (Phila Pa 1976) 34:1923–1928. doi:10.1097/BRS.0b013e3181a9d28e

12. Furlan AD, Pennick V, Bombardier C, van Tulder M (2009) 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. Spine (Phila Pa 1976) 34:1929–1941. doi:10.1097/BRS.0b013e3181b1c99f

13. Kim KT, Lee SH, Lee YH, Bae SC, Suk KS (2006) Clinical outcomes of 3 fusion methods through the posterior approach in the lumbar spine. Spine (Phila Pa 1976) 31(1351):1357. doi:10.1097/01.brs.0000218635.14571.55

14. Furlan AD, Pennick V, Bombardier C, van Tulder M (2009) 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. Spine (Phila Pa 1976) 34:1929–1941. doi:10.1097/BRS.0b013e3181b1c99f

15. van Tulder M, Furlan A, Bombardier C, Bouter L (2003) Updated method guidelines for systematic reviews in the Cochrane collaboration back review group. Spine (Phila Pa 1976) 28:1290–1299. doi:10.1097/01.BRS.0000065484.95996.AF

16. Cowley DE (1995) Prostheses for primary total hip replacement. A critical appraisal of the literature. Int J Technol Assess Health Care 11:770–778

17. Ostelo RW, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, Bouter LM, de Vet HC (2008) Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine (Phila Pa 1976) 33:90–94. doi:10.1097/BRS.0b013e31815e3a10

18. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D, Hill S, Jaeschke R, Leng G, Liberati A, Magrini N, Mason J, Middleton P, Mrukowicz J, O'Connell D, Oxman AD, Phillips B, Schunemann HJ, Edejer T, Varonen H, Vist GE, Williams JJ, Zaza S (2004) Grading quality of evidence and strength of recommendations. BMJ 328:1490. doi:10.1136/bmj.328.7454.1490

19. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, Montori V, Akl EA, Djulbegovic B, Falck-Ytter Y, Norris SL, Williams JJ, Atkins D, Meerpohl J, Schunemann HJ (2011) GRADE guidelines: 4. Rating the quality of evidence–study limitations (risk of bias). J Clin Epidemiol 64:407–415. doi:10.1016/j.jclinepi.2010.07.017

20. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Glasziou P, Jaeschke R, Akl EA, Norris S, Vist G, Dahm P, Shukla VK, Higgins J, Falck-Ytter Y, Schunemann HJ (2011) GRADE guidelines: 7. Rating the quality of evidence–inconsistency. J Clin Epidemiol 64:1294–1302. doi:10.1016/j.jclinepi.2011.03.017

21. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Falck-Ytter Y, Jaeschke R, Vist G, Akl EA, Post PN, Norris S, Meerpohl J, Shukla VK, Nasser M, Schunemann HJ (2011) GRADE guidelines: 8. Rating the quality of evidence–indirectness. J Clin Epidemiol 64:1303–1310. doi:10.1016/j.jclinepi.2011.04.014

22. Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Rind D, Devereaux PJ, Montori VM, Freyschuss B, Vist G, Jaeschke R, Williams JJ, Murad MH, Sinclair D, Falck-Ytter Y, Meerpohl J, Whittington C, Thorlund K, Andrews J, Schunemann HJ (2011) GRADE guidelines 6. Rating the quality of evidence–imprecision. J Clin Epidemiol 64:1283–1293. doi:10.1016/j.jclinepi.2011.01.012

23. Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, Alonso-Coello P, Djulbegovic B, Atkins D, Falck-Ytter Y, Williams JJ, Meerpohl J, Norris SL, Akl EA, Schunemann HJ (2011) GRADE guidelines: 5. Rating the quality of evidence–publication bias. J Clin Epidemiol 64:1277–1282. doi:10.1016/j.jclinepi.2011.01.011

24. Guyatt GH, Oxman AD, Sultan S, Glasziou P, Akl EA, Alonso-Coello P, Atkins D, Kunz R, Brozek J, Montori V, Jaeschke R, Rind D, Dahm P, Meerpohl J, Vist G, Berliner E, Norris S, Falck-Ytter Y, Murad MH, Schunemann HJ (2011) GRADE guidelines: 9. Rating up the quality of evidence. J Clin Epidemiol 64:1311–1316. doi:10.1016/j.jclinepi.2011.06.004

25. Altaf F, Jalgaonkar A, Raman SA (2011) Prospective study of posterior lumbar interbody fusion and posterolateral fusion for degenerative spondylolisthesis. Spine Journal 11:112S–113S

26. Lolli F, Barbanti Brodano G, Di Silvestre M, Martikos K, Vommaro F, Maredi E, Greggi T (2011) Posterolateral instrumented fusion (PLF) versus posterior lumbar interbody fusion (PLIF) in the treatment of low-grade adult isthmic spondylolisthesis. J Orthop Traumatol 12:S42–S43

27. Musluman AM, Yilmaz A, Cansever T, Cavusoglu H, Colak I, Genc HA, Aydin Y (2011) Posterior lumbar interbody fusion versus posterolateral fusion with instrumentation in the treatment of low-grade isthmic spondylolisthesis: midterm clinical outcomes. J Neurosurg Spine 14:488–496. doi:10.3171/2010.11.SPINE10281

28. Cheng L, Nie L, Zhang L (2009) Posterior lumbar interbody fusion versus posterolateral fusion in spondylolisthesis: a prospective controlled study in the Han nationality. Int Orthop 33:1043–1047. doi:10.1007/s00264-008-0588-x

29. Farrokhi MR, Rahmanian A, Masoudi MS (2012) Posterolateral versus posterior interbody fusion in isthmic spondylolisthesis. J Neurotraum 29:1567–1573

30. Inamdar DN, Alagappan M, Shyam L, Devadoss S, Devadoss A (2006) Posterior lumbar interbody fusion versus intertransverse fusion in the treatment of lumbar spondylolisthesis. J Orthop Surg (Hong Kong) 14:21–26

31. Dantas FL, Prandini MN, Ferreira MA (2007) Comparison between posterior lumbar fusion with pedicle screws and posterior lumbar interbody fusion with pedicle screws in adult spondylolisthesis. Arq Neuropsiquiatr 65:764–770

32. Dehoux E, Fourati E, Madi K, Reddy B, Segal P (2004) Posterolateral versus interbody fusion in isthmic spondylolisthesis: functional results in 52 cases with a minimum follow-up of 6 years. Acta Orthop Belg 70:578–582

33. La Rosa G, Cacciola F, Conti A, Cardali S, La Torre D, Gambadauro NM, Tomasello F (2001) Posterior fusion compared with posterior interbody fusion in segmental spinal fixation for adult spondylolisthesis. Neurosurg Focus 10:E9

34. Cunningham JE, Elling EM, Milton AH, Robertson PA (2013) What is the optimum fusion technique for adult isthmic spondylolisthesis—PLIF or PLF? A long-term prospective cohort comparison study. J Spinal Disord Tech 26:260–267. doi:10.1097/BSD.0b013e3182417103

35. Zhao QH, Tian JW, Wang L, Dong SH, Wu ZK, Wang Z, Jia LS (2009) Posterior fusion versus posterior interbody fusion in segmental spinal fixation for aged spondylolisthesis. Zhonghua Yi Xue Za Zhi 89:1779–1782

36. La Rosa G, Conti A, Cacciola F, Cardali S, La Torre D, Gambadauro NM, Tomasello F (2003) Pedicle screw fixation for isthmic spondylolisthesis: does posterior lumbar interbody fusion improve outcome over posterolateral fusion? J Neurosurg 99:143–150. doi:10.3171/spi.2003.99.2.0143

37. DiPaola CP, Molinari RW (2008) Posterior lumbar interbody fusion. J Am Acad Orthop Surg 16:130–139

38. Christensen FB, Hansen ES, Eiskjaer SP, Hoy K, Helmig P, Neumann P, Niedermann B, Bunger CE (2002) Circumferential lumbar spinal fusion with Brantigan cage versus posterolateral fusion with titanium Cotrel-Dubousset instrumentation: a prospective, randomized clinical study of 146 patients. Spine (Phila Pa 1976) 27:2674–2683. doi:10.1097/01.BRS.0000035264.64371.87

39. Wetzel FT, Brustein M, Phillips FM, Trott S (1999) Hardware failure in an unconstrained lumbar pedicle screw system. A 2-year follow-up study. Spine (Phila Pa 1976) 24:1138–1143

40. Kim KT, Lee SH, Lee YH, Bae SC, Suk KS (2006) Clinical outcomes of 3 fusion methods through the posterior approach in the lumbar spine. Spine (Phila Pa 1976) 31(1351):1357. doi: 10.1097/01.brs.0000218635.14571.55

41. Pichelmann MA, Lenke LG, Bridwell KH, Good CR, O'Leary PT, Sides BA (2010) Revision rates following primary adult spinal deformity surgery: six hundred forty-three consecutive patients followed-up to twenty-two years postoperative. Spine (Phila Pa 1976) 35:219–226. doi:10.1097/BRS.0b013e3181c91180

42. Raizman NM, O'Brien JR, Poehling-Monaghan KL, Yu WD (2009) Pseudarthrosis of the spine. J Am Acad Orthop Surg 17:494–503

43. Yu CH, Wang CT, Chen PQ (2008) Instrumented posterior lumbar interbody fusion in adult spondylolisthesis. Clin Orthop Relat Res 466:3034–3043. doi:10.1007/s11999-008-0511-1

44. Wetzel FT, Brustein M, Phillips FM, Trott S (1999) Hardware failure in an unconstrained lumbar pedicle screw system. A 2-year follow-up study. Spine (Phila Pa 1976) 24:1138–1143

45. Ekman P, Moller H, Tullberg T, Neumann P, Hedlund R (2007) Posterior lumbar interbody fusion versus posterolateral fusion in adult isthmic spondylolisthesis. Spine (Phila Pa 1976) 32:2178–2183. doi:10.1097/BRS.0b013e31814b1bd8

46. Lidar Z, Beaumont A, Lifshutz J, Maiman DJ (2005) Clinical and radiological relationship between posterior lumbar interbody fusion and posterolateral lumbar fusion. Surg Neurol 64(303–308):308. doi:10.1016/j.surneu.2005.03.025

47. Jacobs WC, Vreeling A, De Kleuver M (2006) Fusion for low-grade adult isthmic spondylolisthesis: a systematic review of the literature. Eur Spine J 15:391–402. doi:10.1007/s00586-005-1021-4

48. Andersen T, Christensen FB, Hansen ES, Bunger C (2003) Pain 5 years after instrumented and non-instrumented posterolateral lumbar spinal fusion. Eur Spine J 12:393–399. doi:10.1007/s00586-003-0547-6

49. Schnee CL, Freese A, Ansell LV (1997) Outcome analysis for adults with spondylolisthesis treated with posterolateral fusion and transpedicular screw fixation. J Neurosurg 86:56–63. doi:10.3171/jns.1997.86.1.0056

50. Fritzell P, Hagg O, Wessberg P, Nordwall A (2002) Chronic low back pain and fusion: a comparison of three surgical techniques: a prospective multicenter randomized study from the Swedish lumbar spine study group. Spine (Phila Pa 1976) 27:1131–1141

51. Wright JG (1996) The minimal important difference: who's to say what is important? J Clin Epidemiol 49:1221–1222

52. Vamvanij V, Fredrickson BE, Thorpe JM, Stadnick ME, Yuan HA (1998) Surgical treatment of internal disc disruption: an outcome study of four fusion techniques. J Spinal Disord 11:375–382

53. Vlaanderen J, Vermeulen R, Heederik D, Kromhout H (2008) Guidelines to evaluate human observational studies for quantitative risk assessment. Environ Health Perspect 116:1700–1705. doi:10.1289/ehp.11530

54. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB (2000) Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of observational studies in epidemiology (MOOSE) group. JAMA 283:2008–2012

55. Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S, Guyatt GH (2011) GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 64:401–406. doi:10.1016/j.jclinepi.2010.07.015

56. Egger M, Ebrahim S, Smith GD (2002) Where now for meta-analysis? Int J Epidemiol 31:1–5

57. Cunningham J, Robertson P (2011) Long-term outcomes following lumbar spine fusion for adult isthmic spondylolisthesis: a comparison of PLIF versus PLF. Spine J 11:135S