Decompression alone or decompression and fusion in degenerative lumbar spondylolisthesis

Fei-Long Wei,a,1 Cheng-Pei Zhou,a,1 Quan-You Gao,a,1 Ming-Rui Du,a Hao-Ran Gao,a Kai-Long Zhu,a Tian Li,b** Ji-Xian Qian,a* and Xiao-Dong Yan,a*

aDepartment of Orthopaedics, Tangdu Hospital, Fourth Military Medical University, 710038, Xi’an, China
bSchool of Basic Medicine, Fourth Military Medical University, 710032, Xi’an, China

Summary

Background Clinically, there are substantive practice variations in surgical management of degenerative lumbar spondylolisthesis. We aimed at evaluating whether decompression alone outcomes for patients with degenerative lumbar spondylolisthesis are comparable to those of decompression with fusion.

Methods In this meta-analysis, the Embase, PubMed, and Cochrane Library databases were searched from inception to February 16th, 2022. Randomised controlled trials (RCTs) and cohort studies comparing decompression alone with decompression and fusion for patients with degenerative lumbar spondylolisthesis were included in this study. There were no language limitations. Odds ratio (OR), mean difference (MD) and 95% confidence interval (CI) were used to report results in the random-effects model. Main outcomes included Oswestry disability index (ODI), pain, clinical satisfaction, complication and reoperation rates. The study protocol was published in PROSPERO (CRD42022310645).

Findings Thirty-three studies (6 RCTs and 27 cohort studies) involving 94 953 participants were included. Differences in post-operative ODI between decompression alone and decompression with fusion were not significant. A small difference for back (MD, 0.13; [95% CI, 0.08 to 0.18]; I²:48.35%) and leg pain (MD, 0.30; [95% CI, 0.09 to 0.51]; I²:48.35%) was observed on the 3rd post-operative month. The results did not reveal significant differences in leg pain and back pain between decompression alone and fusion groups on the 6th, 12th, and 24th post-operative months. Difference in clinical satisfaction between decompression alone and decompression with fusion were not significant from RCTs (OR, 0.26; [95% CI, 0.03 to 1.92]; I²:83.27%). Complications (OR, 1.54; [95% CI, 1.16 to 2.05]; I²:48.88%), operation time (MD, 83.39; [95% CI, 55.93 to 110.85]; I²:98.75%), intra-operative blood loss (MD, 264.58; [95% CI, 174.99 to 354.16]; I²:95.61%) and length of hospital stay (MD, 2.85; [95% CI, 1.60 to 4.10]; I²:99.49%) were higher with fusion.

Interpretation Clinical effectiveness of decompression alone was comparable to that of decompression with fusion for degenerative lumbar spondylolisthesis. Decompression alone is recommended for patients with degenerative lumbar spondylolisthesis.

Funding This work was supported by grants from the National Natural Science Foundation of China (No. 81871818), Tangdu Hospital Seed Talent Program (Fei-Long Wei), Natural Science Basic Research Plan in Shaanxi Province of China (No.2019JM-265) and Social Talent Fund of Tangdu Hospital (No.2021SHRC034).

Copyright © 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Keywords: Fusion; Decompression; Degenerative lumbar spondylolisthesis; Outcomes
Research in context

Evidence before this study

Clinically, there are substantive practice variations in surgical management of degenerative lumbar spondylolisthesis. We searched the Cochrane and PROSPERO databases using terms, including “decompression”, “fusion”, “spondylolisthesis” to identify previous meta-analyses on this topic. Previous meta-analyses have reported conflicting results with regarding decompression alone or with fusion for degenerative lumbar spondylolisthesis.

Added value of this study

In this updated meta-analysis, we found that the clinical effectiveness of decompression alone was comparable to that of decompression and fusion for degenerative lumbar spondylolisthesis. Complications rate, operation time, intra-operative blood loss and length of hospital stay were higher with fusion.

Implications of all the available evidence

Our findings do not support routine applications of decompression and fusion for degenerative lumbar spondylolisthesis. Decompression alone is recommended for patients with degenerative lumbar spondylolisthesis.

Methods

This meta-analysis was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guidelines. The study protocol was published in PROSPERO (CRD42022310645).

Search strategy and selection criteria

The Cochrane and PROSPERO databases were independently searched by two reviewers (F.-L. W. and C.-P. Z.) to avoid missing relevant studies. The Embase, PubMed, and Cochrane Library databases were searched from inception to February 16th, 2022. There were no language limitations in the search. The search strategy is presented in eTable 1 in the Supplement. After preliminary screening of titles or abstracts, two independent reviewers (F.-L. W. and C.-P. Z.) evaluated the related publications.

Studies were screened based on the PICOS criteria, as documented in eTable 2 in the Supplement. The inclusion criteria were the following: 1) The study design was RCT or non-RCT comparing decompression alone (including open decompression and minimally invasive decompression) with decompression plus fusion; 2) Adult patients over the age of 18 with degenerative lumbar spondylolisthesis; 3) The study reported at least one outcome; 4) Each group had at least 5 patients. The exclusion criteria were as follows: 1) Non-controlled; 2) Patients suffering from trauma, spinal tumors, or infection or with isthmic spondylolisthesis; 3) Surgery was performed using an anterior approach; 4) Studies that were repeatedly published or had qualitative outcomes; 5) Quasi-experimental studies, crossover, and observational studies.

Data extraction and outcomes

Data extraction from included articles was performed by two independent reviewers (F.-L. W. and C.-P. Z.). Extracted data included the characteristics of investigators, study types, surgical methods, participant characteristics, Grade (s) of degenerative lumbar spondylolisthesis and main outcomes. Primary outcomes were Oswestry disability index (ODI), pain (visual analogue scores (VAS) or numerical rating scale (NRS)), clinical satisfaction, complication rates and reoperation alone is non-inferior to decompression with instrumented fusion. However, during a 4-year follow-up, a randomised controlled trial (RCT) showed that reoperation occurred frequently in the decompression alone group (34%), relative to the fusion group (14%). Therefore, we evaluated whether clinical outcomes from decompression alone are comparable to decompression with fusion outcomes in patients with degenerative lumbar spondylolisthesis.

Introduction

Degenerative lumbar spondylolisthesis, in which one vertebral body slips forward relative to the vertebral body below with an intact neural arch, is common in people aged over 60 years. This condition is associated with leg and back pain as well as functional limitations. It disproportionately affects women, especially black women, with a male to female ratio of about 1:6. Spondylolisthesis commonly occurs at the L4-L5 level, and rarely exceeds 30% of the vertebral body width. Due to the aging global population, degenerative lumbar spondylolisthesis is an important cause of disability. Surgical therapy is recommended for patients who fail conservative treatment.

Surgical options include decompression alone or decompression with fusion. In the United States, fusion rates more than doubled from 2005 to 2014, with degenerative lumbar spondylolisthesis accounting for majority of fusion rates. Moreover, hospitalization costs for lumbar internal fixation and fusion were estimated at $13 billion in 2011, higher than any other surgical procedure in the U.S. The need for fusion surgery remains controversial.

Two meta-analyses have reported conflicting outcomes with regards to decompression alone or with fusion for degenerative lumbar spondylolisthesis. In 2020, real-world evidence suggested that decompression
rates. Secondary outcomes were blood loss, operative time, and hospital stay.

Risk of bias assessment
The Cochrane Collaboration’s tool\textsuperscript{17} was used by two reviewers (F.-L. W. and C.-P. Z.) to independently evaluate the included RCTs for potential bias. The detailed information of the tool for assessing the risk of bias is provided in eTable 3 in the Supplement. Overall risk of bias was divided into “high risk”, “low risk”, or “unclear risk”. Qualities of the included cohort studies were evaluated by the Newcastle-Ottawa Quality Assessment Scale (NOS).\textsuperscript{18} A high-quality study had a NOS score >6. Disagreements between the two investigators were resolved via discussions involving a third investigator (X.-D. Y.).

Data analysis
STATA 16.0 (Stata Corp, College Station, TX, USA) was used for statistical analyses. Data pooling was done using a random-effects model. Dichotomous variables were evaluated by odds ratio (OR) with 95% confidence interval (CI). Mean difference (MD) with 95% CI was used to weigh the effect sizes for continuous outcomes. Effect sizes were assessed by a forest plot. Weights of the included studies were dependent on the value of the event in the decompression group, the event in the decompression with fusion group, and the size of the entire sample. \( p \leq 0.05 \) denoted significant differences. Statistical heterogeneity among summary data were evaluated using \( I^2 \) statistic. If the test showed \( I^2 > 50\% \), data had a high heterogeneity. For primary outcomes (ODI, back pain and leg pain) were reported on the 3rd, 6th, 12th and 24th post-operative months.
For studies reporting results at multiple time points, the data reported at the time points closest to 3rd, 6th, 12th and 24th post-operative months will be included in the primary analysis. Reoperation rate was divided into short-term (<4 years) and long-term (≥4 years). Then subgroup analysis was conducted to explore heterogeneity depending on trial types. If the results are inconsistent, we will interpret them according to the results derived from RCTs. In addition, sensitivity analysis was conducted by eliminating low-quality and older studies.

Ethics statement
Ethical approval for this study is not applicable since the data utilized were collected from previously published research in the literature. All the included studies in this study had received ethical approval prior to data collection.

Role of the funding source
The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. All authors had full access to all the data in the study, and accept responsibility to submit for publication.

Results

Literature search results
Our literature search, curation, and analysis, did not reveal any duplicate meta-analysis themes in Cochrane and PROSPERO. The literature search process is summarized in Figure 1. Through database searches and manual searches of reference lists for relevant literature reviews, 3014 study records were identified. After removing duplicates and screening the titles and abstracts for the remaining articles, 81 full-text articles were evaluated. Ultimately, 33 studies (6 RCTs7,8,19−22 and 27 cohort studies11,23−46) involving 94,953 participants were included in this study.

Study characteristics
The characteristics of the included studies are shown in eTable 4 in the Supplement. The studies had been performed in North America, Europe, and Asia with publications being performed between 1991 and 2022. Based on our defined outcomes, 13 studies reported on ODI outcomes7,8,11,19,20,23,24,26,27,31,34−36; 14 reported on back pain outcomes8,11,19,20,23,24,26,27,31,32,34−36; 14 reported on leg pain outcomes8,11,19,20,23,24,26,27,31,34−36; 18 reported on complication outcomes7,8,11,19,20,23,24,26,27,31,32,34−36; 22 reported on reoperation outcomes8,11,19,20,23,24,26,27,31,32,34−36; 9 reported on intra-operative blood loss outcomes8,19,20,24,26,27,31,32; while 12 reported on length of hospital stay outcomes8,19,20,23,24,26,27,31,32. A summary of the risk of bias assessment of the RCTs is displayed in eFigure 1 and 2 in the Supplement. The risks of bias of the included cohort studies are displayed in eTable 5 in the Supplement.

Primary outcomes

ODI. Pooled analysis of 12 studies did not reveal significant differences in ODI between decompression alone and fusion groups on the 3rd, 6th, 12th, and 24th post-operative months (Figure 2). Decompression alone was not inferior to decompression with fusion in improving patient dysfunction. For studies reporting on ODI on the 3rd and 24th post-operative months, heterogeneity was more than 50% (Figure 2). A funnel plot (eFigure 3 in the Supplement) showed deviations in publications. Subgroup analysis revealed that the trial types did not have any effects on ODI on the 3rd, 12th and 24th post-operative months (eFigure 4, eFigure 5 and eFigure 6 in the Supplement). Since the data bias of Chan 2019a26 and Chan 2019b27 is large, we added sensitivity analysis by eliminating these two trials. These studies showed that fusion surgery was associated with superior ODI. The result was consistent with primary results (eFigure 7 in the Supplement). But the heterogeneity is much reduced.

Back pain

On the 3rd post-operative month, pooled analysis of 14 studies showed significantly higher back pain in the decompression group, relative to the fusion group (MD, 0.13; [95% CI, 0.08 to 0.18]; Figure 3). However, fusion had no meaningful impact on back pain on the 6th, 12th and 24th post-operative months (Figure 3). For studies reporting back pain on the 12th and 24th post-operative months, there was more than 50% heterogeneity (Figure 3). A funnel plot (eFigure 8 in the Supplement) showed deviations in publications. Subgroup analysis revealed that on the 3rd post-operative month, trial types had effects on back pain (eFigure 9 in the Supplement). But subgroup analysis revealed that on the 24th post-operative month, trial types had no effects on back pain (eFigure 10 in the Supplement). Subgroup analyses were not performed because the 12-month data were all from observational cohort studies. Since the data bias of Chan 2019a26 and Chan 2019b27 is large, we added sensitivity analysis by eliminating these two trials. The study77 showed that fusion surgery was associated with superior back pain. The sensitivity analysis was consistent with primary results (eFigure 11 in the Supplement). But the heterogeneity is much reduced.
Figure 2. Forest plot for comparisons of ODI between decompression and decompression with fusion groups. SD: standard deviation; CI: confidence interval; ODI: Oswestry Disability Index; B<L: Back pain<Leg pain; B>L: Back pain>Leg pain.
### Figure 3.

Forest plot for comparisons of back pain between decompression and decompression with fusion groups. SD: standard deviation; CI: confidence interval. EF+: Effusion; EF-: No effusion; B<L: Back pain<Leg pain; B>L: Back pain>Leg pain.

| Study            | Decompression | Fusion          | Mean Diff. with 95% CI | Weight (%) |
|------------------|---------------|-----------------|------------------------|------------|
|                  | N  | Mean | SD | N  | Mean | SD |                     |             |
| **3 Months**     |    |      |    |    |      |    |                     |             |
| Austevoll 2016   | 212 | 3.6  | 2.7 | 202 | 3.1  | 2.3 |                     | 0.50 [ 0.02, 0.98] | 4.55       |
| Austevoll 2020   | 285 | -3.3 | 2.9 | 285 | -3.6 | 2.9 |                     | 0.30 [ -0.18, 0.78] | 4.56       |
| Austevoll 2021   | 130 | 3.38 | .1925 | 125 | 3.25 | .195 |                     | 0.13 [ 0.08, 0.18] | 4.97       |
| Hua 2021         | 24  | 2.3  | .5  | 36  | 2.4  | .8  |                     | -0.10 [ -0.46, 0.26] | 4.73       |
| **6 Months**     |    |      |    |    |      |    |                     |             |
| Staartjes 2018   | 35  | 2.3  | 2  | 41  | 2.8  | 2 |                     | -0.50 [ -1.40, 0.40] | 3.76       |
| **12 Months**    |    |      |    |    |      |    |                     |             |
| Kleinsteuck 2012 | 54  | 2.4  | 3.69 | 156 | 2.4  | 2.9 |                     | 0.00 [ -0.97, 0.97] | 3.63       |
| Sigmundsson 2015 (B<L) | 100 | .86 | 2.98 | 213 | 2.68 | 2.98 |                     | -1.82 [ -2.53, -1.11] | 4.15       |
| Sigmundsson 2015 (B>L) | 97  | 2.95 | 3.01 | 207 | 3.91 | 3.26 |                     | -0.96 [ -1.73, -0.19] | 4.03       |
| Alvin 2016       | 35  | 3.64 | 2.42 | 75  | 2.5  | 2.45 |                     | 1.10 [ -0.01, 2.21] | 3.35       |
| Austevoll 2016   | 216 | 3.9  | 2.9 | 224 | 3.3  | 2.6 |                     | 0.60 [ 0.09, 1.11] | 4.50       |
| Staartjes 2018   | 30  | 2.7  | 2.3 | 51  | 2.7  | 2.6 |                     | 0.00 [ -1.12, 1.12] | 3.31       |
| Chan 2019b       | 84  | -3.1 | 3.6 | 342 | -4.1 | 3 |                     | 1.00 [ 0.25, 1.75] | 4.07       |
| Austevoll 2020   | 285 | 3.8  | 2.9 | 285 | 3.3  | 2.6 |                     | 0.50 [ 0.05, 0.95] | 4.60       |
| Hua 2021         | 24  | 1.8  | .5 | 36  | 1.9  | .7 |                     | -0.10 [ -0.42, 0.22] | 4.77       |
| **24 Months**    |    |      |    |    |      |    |                     |             |
| Park 2012        | 20  | 1.2  | 2.2 | 25  | 2.4  | 1.88 |                     | -1.20 [ -2.39, -0.01] | 3.18       |
| Forsth 2013      | 655 | 3.5  | 3.264 | 651 | 3.2  | 2.604 |                     | 0.30 [ -0.02, 0.62] | 4.78       |
| Sigmundsson 2015 (B<L) | 73 | .81 | 2.82 | 130 | 2.23 | 2.82 |                     | -1.42 [ -2.23, -0.61] | 3.95       |
| Sigmundsson 2015 (B>L) | 76 | 2.73 | 3.12 | 148 | 3.36 | 2.74 |                     | -0.63 [ -1.42, 0.16] | 3.98       |
| Lattig 2015 (EF+) | 33 | 1.9 | 3.6 | 69 | 1.7 | 3 |                     | 0.20 [ -1.13, 1.53] | 2.93       |
| Lattig 2015 (EF-) | 18 | 1.1 | 3 | 20 | 2.9 | 2.6 |                     | -1.80 [ -3.58, -0.02] | 2.21       |
| Forsth 2016      | 66  | 2.6  | 2.5 | 67  | 3.6  | 2.9 |                     | -1.00 [ -1.92, -0.08] | 3.72       |
| Staartjes 2018   | 33  | 2.1  | 2.1 | 51  | 2.8  | 2.6 |                     | -0.80 [ -1.86, 0.26] | 3.44       |
| Chan 2019a       | 71  | -1.5 | 4.4 | 72 | -4.7 | 3.2 |                     | 3.20 [ 1.94, 4.46] | 3.05       |
| Austevoll 2021   | 119 | 3.33 | .205 | 121 | 3.67 | .2075 |                     | -0.34 [ -0.39, -0.29] | 4.97       |
| Hua 2021         | 24  | 1.7  | .5 | 36  | 1.8  | .6 |                     | -0.10 [ -0.39, 0.19] | 4.81       |
| **Overall**      |    |      |    |    |      |    |                     |             |
|                  |    |      |    |    |      |    |                     |             |
|                  |    |      |    |    |      |    |                     | -0.10 [ -0.45, 0.25] |             |
|                  |    |      |    |    |      |    |                     |             |

**Random-effects REML model**

Favours decompression Favours fusion

---

**Figure 3.** Forest plot for comparisons of back pain between decompression and decompression with fusion groups. SD: standard deviation; CI: confidence interval. EF+: Effusion; EF-: No effusion; B<L: Back pain<Leg pain; B>L: Back pain>Leg pain.
Leg pain
On the 3rd post-operative month, pooled analysis of 14 studies showed significantly higher leg pain in the decompression group, relative to the fusion group (MD, 0.30; [95% CI, 0.09 to 0.51]; Figure 4). Studies reporting on leg pain on the 12th and 24th post-operative months, fusion had no meaningful impact on back pain (Figure 4). Studies reporting on leg pain on the 12th and 24th post-operative months had more than 50% heterogeneity (Figure 4). A funnel plot (eFigure 12 in the Supplement) showed deviations in publications. Subgroup analyses revealed that on the 3rd and 12th post-operative months, trial types did not have effects on leg pain (eFigure 13 and eFigure 14 in the Supplement). On the 24th post-operative month (eFigure 15 in the Supplement), pooled analysis of 3 RCTs showed that fusion had meaningful impact on leg pain (MD, −0.27; [95% CI, −0.32 to −0.21]). Due to the bias of Chan 2019a,26 and Chan 2019b,27 we added sensitivity analysis by eliminating these two trials. The result was consistent with primary results (eFigure 16 in the Supplement).

Clinical satisfaction
Pooled analysis of 12 studies showed that decompression alone was associated with a low clinical satisfaction rate (OR, 0.57; [95% CI, 0.35 to 0.91]; Figure 5A). More than 50% heterogeneity was found in studies reporting on clinical satisfaction (Figure 5A). More than two types of scoring systems were used to assess clinical satisfaction (eTable 4). A funnel plot (eFigure 17 in the Supplement) showed deviations in publications. Subgroup analysis revealed that difference in clinical satisfaction between decompression alone and decompression with fusion was not significant from RCTs (OR, 0.26; [95% CI, 0.03 to 1.92]) (Figure 5A). Sensitivity analysis by eliminating low-quality and older studies21,48 showed that fusion had no meaningful effects on clinical satisfaction which was consistent with the result derived from RCTs (eFigure 18 in the Supplement).

Complications
Pooled analysis of 18 studies showed that decompression alone was associated with low complication rates (OR, 0.65; [95% CI, 0.49 to 0.86]; Figure 3B). There was a less than 50% heterogeneity for studies reporting on complication rates (Figure 3B). A funnel plot (eFigure 19 in the Supplement) showed deviations in publications. Subgroup analysis revealed that trial types had no effects on complication rates (eFigure 20 in the Supplement). Sensitivity analysis by eliminating low-quality and older studies21,48 was consistent with primary results (eFigure 21 in the Supplement).

Reoperation
Pooled analysis of 22 studies revealed that fusion had no meaningful effects on reoperation rates (Figure 6). Short-term (<4 years) reoperation rates in the decompression group and fusion group were 4.78% and 3.43%, respectively. Long-term (≥4 years) reoperation rates were markedly increased in both groups (decompression group, 16.0%; fusion group, 17.43%) without significant differences. Less than 50% heterogeneity was found in studies reporting on reoperation (Figure 6). A funnel plot (eFigure 22 in the Supplement) showed deviations in publications. Subgroup analysis revealed that trial types had no effects on reoperation (eFigure 23 and eFigure 24 in the Supplement). Sensitivity analysis by eliminating low-quality and older studies21 was consistent with primary results (eFigure 25 in the Supplement).

Secondary outcomes

Operation time. Pooled analysis of 11 studies showed that decompression alone was associated with less operation time (MD, −83.39; [95% CI, −110.85 to −55.93]; eFigure 26 in the Supplement). More than 50% heterogeneity was found in studies reporting on operation time (eFigure 26 in the Supplement). Subgroup analysis revealed that trial types did not have effects on operation time (eFigure 27 in the Supplement).

Intra-operative blood loss
Pooled analysis of 9 studies showed that decompression alone was associated with less intra-operative blood loss (MD, −264.58; [95% CI, −354.16 to −174.99]; eFigure 28 in the Supplement). Studies reporting on intra-operative blood loss had more than 50% heterogeneity (eFigure 28 in the Supplement). Subgroup analysis revealed that trial types had no effects on intra-operative blood loss (eFigure 29 in the Supplement).

Length of hospital stay
Pooled analysis of 12 studies showed that decompression alone was associated with a shorter length of hospital stay (MD, −2.85; [95% CI, −4.10 to −1.60]; eFigure 30 in the Supplement). More than 50% heterogeneity was found in studies reporting on length of hospital stay (eFigure 30 in the Supplement). Subgroup analysis revealed that trial types had no effects on length of hospital stay (eFigure 31 in the Supplement).

Discussion
Lumbar interbody fusion has attracted the attention of many spine surgeons to treat a range of spinal disorders since it was first described by Briggs and Milligan in 1944.49 Evidence for the use of fusion in degenerative lumbar spondylolisthesis is insufficient. A small difference for back and leg pain was observed on the 3rd post-
**Figure 4.** Forest plot for comparisons of leg pain between decompression and decompression with fusion groups. SD: standard deviation; CI: confidence interval. EF+: Effusion; EF−: No effusion; B<L: Back pain<Leg pain; B>L: Back pain>Leg pain.
operative month. And these effects disappeared after 3 months. Fusion was associated with longer operative time, more intra-operative blood loss, longer hospital stay and more complications. Therefore, our findings do not support routine use of decompression with fusion in patients with degenerative lumbar spondylolisthesis.

Hospital costs for elective lumbar degenerative fusion exceeded $10 billion in 2015, the highest total cost of any surgical procedure in the United States.\textsuperscript{30} Lumbar fusion surgery remains controversial for many indications.\textsuperscript{6} Previous RCT by Ghogawala et al in 2016\textsuperscript{7} and clinical guidelines\textsuperscript{3,51} supported that spinal fusion may lead to better clinical outcomes than decompression alone for degenerative lumbar spondylolisthesis. Two previous well-designed RCTs\textsuperscript{8,20} and a systematic review\textsuperscript{52} have challenged the widespread use of fusion in the surgical treatment of degenerative lumbar spondylolisthesis. Whether fusion is needed is still a matter of debate. Many surgeons may consider slippage and dynamic instability at the level of spondylolisthesis to be better treated with fusion.\textsuperscript{3} Instability, or the degree of spondylolisthesis in patients with degenerative lumbar spondylolisthesis may be aggravated by decompression alone.\textsuperscript{3,53,54} We incorporated recent RCTs\textsuperscript{19,20} and real-world data\textsuperscript{11} in this study and found significant differences in clinical satisfaction rates between the two groups (OR, 0.57; [95% CI, 0.35 to

| Study        | Decompression Fusion | Odds Ratio with 95% CI | Weight (%) |
|--------------|----------------------|------------------------|------------|
| Yone 1996    | Yes: 2 Yes: 5 No: 8 No: 2 | 0.10 [0.01, 0.95] | 3.64       |
| Matsudaira 2005 | Yes: 15 Yes: 3 No: 13 No: 6 | 2.31 [0.48, 11.12] | 6.07       |
| Kim 2012     | Yes: 46 Yes: 11 No: 49 No: 9 | 0.77 [0.29, 2.02] | 9.99       |
| Kleinstueck 2012 | Yes: 38 Yes: 16 No: 125 No: 20 | 0.38 [0.18, 0.81] | 11.85      |
| Park 2012    | Yes: 13 Yes: 7 No: 14 No: 11 | 1.46 [0.43, 4.90] | 8.17       |
| Rampersaud 2014 | Yes: 29 Yes: 17 No: 88 No: 45 | 0.87 [0.43, 1.75] | 12.32      |
| Austevoi 2016 | Yes: 135 Yes: 82 No: 158 No: 67 | 0.70 [0.47, 1.04] | 14.86      |
| Chan 2019a   | Yes: 40 Yes: 15 No: 55 No: 4 | 0.19 [0.06, 0.63] | 8.41       |
| Hua 2021     | Yes: 23 Yes: 1 No: 35 No: 1 | 0.66 [0.04, 11.04] | 2.53       |
| Heterogeneity: $\chi^2 = 0.11$, $I^2 = 33.95\%$, $H^2 = 1.51$ | | 0.64 [0.43, 0.95] |
| Test of $\theta = \theta_0$: $Q(8) = 13.80$, $p = 0.09$ | | |

| RCT          | Decompression Fusion | Odds Ratio with 95% CI | Weight (%) |
|--------------|----------------------|------------------------|------------|
| Herkowitz 1991 | Yes: 11 Yes: 14 No: 24 No: 1 | 0.03 [0.00, 0.28] | 3.92       |
| Bridwell 1993 | Yes: 3 Yes: 6 No: 23 No: 11 | 0.24 [0.05, 1.14] | 6.13       |
| Forsth 2016   | Yes: 45 Yes: 21 No: 43 No: 24 | 1.20 [0.58, 2.46] | 12.13      |
| Heterogeneity: $\chi^2 = 2.55$, $I^2 = 83.27\%$, $H^2 = 5.98$ | | 0.26 [0.03, 1.92] |
| Test of $\theta = \theta_0$: $Q(2) = 11.73$, $p = 0.00$ | | |

| Overall      | Decompression Fusion | Odds Ratio with 95% CI | Weight (%) |
|--------------|----------------------|------------------------|------------|
| Heterogeneity: $\chi^2 = 0.38$, $I^2 = 62.19\%$, $H^2 = 2.64$ | | 0.57 [0.35, 0.93] |
| Test of $\theta = \theta_0$: $Q(11) = 25.54$, $p = 0.01$ | | |
| Test of group differences: $Q(1) = 0.76$, $p = 0.38$ | | |

\textbf{Figure 5.} (A). Forest plot for comparisons of clinical satisfaction between decompression and decompression with fusion groups; (B). Forest plot for comparisons of complication rates between decompression and decompression with fusion groups. CI: confidence interval.
However, findings from RCTs were inconsistent with that from cohort studies on clinical satisfaction. In our opinion, for patient-reported outcome measures such as clinical satisfaction, when the results are inconsistent, the results should be derived from RCTs. Therefore, the results tended to be that fusion had no effect on clinical satisfaction (derived from RCTs). And we conducted sensitivity analysis by eliminating low-quality and older studies which was consistent with the results derived from RCTs. This result was consistent with one meta-analysis and previous RCTs which was consistent with the results derived from RCTs. This result was consistent with one meta-analysis and previous RCTs that, relative to surgical decompression alone, decompression with fusion have no better clinical outcomes. However, another study suggested that fusion improves clinical satisfaction and reduces post-operative leg pain. The reason for this difference may be that they did not distinguish between comparative cohort studies and RCTs. In addition, more than two types of scoring systems were used to assess clinical satisfaction, so this conclusion should be cautiously interpreted. Although the fusion group exhibited better pain reduction in the short term, the difference was small and unlikely to be clinically meaningful. The results were consistent with latest RCTs. Complication rates are an important indicator to assess the safety of surgery. Previous meta-analyses concluded that decompression with fusion have similar complication rates. However, current study concluded that the decompression alone group had a lower complication rates, relative to the fusion group. Additionally, subgroup analysis showed that trial types had no effects on the results. Observational data from real-world practice and a recent RCT reported comparable conclusions, with higher complication rate in decompression with fusion group. Another reason that supports the surgeon’s choice of fusion is that
Figure 6. Forest plot for comparisons of reoperation rates between decompression and decompression with fusion groups. CI: confidence interval; Short-term: <4 years; Long-term: ≥4 years.

| Study      | Decompression | Fusion | Odds Ratio with 95% CI | Weight (%) |
|------------|---------------|--------|------------------------|------------|
| Short-term |               |        |                        |            |
| Bridwell 1993 | 0   | 9    | 1                      | 33         | 1.18 [0.04, 31.26] | 0.89 |
| Ghogawala 2004 | 3   | 17   | 0                      | 14         | 5.80 [0.28, 121.72] | 1.02 |
| Matsudaiga 2005 | 0  | 18   | 2                      | 17         | 0.19 [0.01, 4.22]  | 0.98 |
| Rampersaud 2014 | 5   | 41   | 9                      | 16         | 0.22 [0.06, 0.75]  | 4.15 |
| Forsth 2016   | 15  | 53   | 14                     | 53         | 1.07 [0.47, 2.44]  | 6.28 |
| Ulrich 2017   | 9   | 76   | 2                      | 44         | 2.61 [0.54, 12.60] | 3.01 |
| Staartjes 2018 | 4   | 47   | 3                      | 48         | 1.36 [0.29, 6.42]  | 3.08 |
| Chan 2019a    | 10  | 61   | 1                      | 71         | 11.64 [1.45, 93.53]| 1.96 |
| Chan 2019b    | 5   | 79   | 15                     | 327        | 1.38 [0.49, 3.91]  | 5.04 |
| Pieters 2019  | 5   | 484  | 171                    | 5,528      | 0.33 [0.14, 0.82]  | 5.85 |
| Badhiwala 2021 | 18  | 784  | 28                     | 974        | 0.80 [0.44, 1.45]  | 7.75 |
| Austevoll 2021 | 15  | 105  | 11                     | 110        | 1.43 [0.63, 3.25]  | 6.28 |
| Heterogeneity: $\chi^2 = 0.39$, $I^2 = 53.49\%$, $H^2 = 2.15$ |  |        |                        |            | 1.00 [0.59, 1.70]  | 5.04 |
| Test of $\theta = 0$; $Q(11) = 22.24$, $p = 0.02$ |  |        |                        |            |                 |

| Long-term   |               |        |                        |            |
| Kim 2012    | 8   | 49   | 6                      | 52         | 1.41 [0.46, 4.37]  | 4.62 |
| Park 2012   | 1   | 19   | 0                      | 25         | 3.92 [0.15, 101.63]| 0.90 |
| Sato 2015   | 25  | 49   | 13                     | 76         | 2.98 [1.39, 6.38]  | 6.67 |
| Ghogawala 2016 | 12  | 23   | 4                      | 27         | 3.52 [1.00, 12.43] | 4.05 |
| Inui 2016   | 4   | 56   | 11                     | 69         | 0.45 [0.14, 1.48]  | 4.31 |
| Vorhies 2018 | 351 | 1,723| 3,183                  | 14,843     | 0.95 [0.84, 1.07]  | 10.42 |
| Kimura 2019 | 2   | 26   | 4                      | 46         | 0.88 [0.15, 5.16]  | 2.55 |
| Kuo 2019    | 17  | 147  | 75                     | 362        | 0.56 [0.32, 0.98]  | 8.03 |
| Joelson 2021 | 80  | 517  | 216                    | 1,122      | 0.80 [0.61, 1.06]  | 9.81 |
| Inose 2022  | 2   | 27   | 3                      | 28         | 0.69 [0.11, 4.47]  | 2.34 |
| Heterogeneity: $\chi^2 = 0.25$, $I^2 = 78.69\%$, $H^2 = 4.69$ |  |        |                        |            | 1.06 [0.69, 1.64]  | 5.04 |
| Test of $\theta = 0$; $Q(9) = 20.48$, $p = 0.02$ |  |        |                        |            |                 |

| Overall     |               |        |                        |            |
|             |               |        |                        |            |
|             |               |        |                        |            |
|             |               |        |                        |            |
|             |               |        |                        |            |
|             |               |        |                        |            |
|             |               |        |                        |            |
| Heterogeneity: $\chi^2 = 0.26$, $I^2 = 70.11\%$, $H^2 = 3.35$ |  |        |                        |            | 1.02 [0.74, 1.41]  | 5.04 |
| Test of $\theta = 0$; $Q(21) = 42.76$, $p = 0.00$ |  |        |                        |            |                 |
| Test of group differences: $Q(1) = 0.03$, $p = 0.86$ |  |        |                        |            |                 |
decompression alone will lead to instability, which in turn will result in a higher reoperation rate.\textsuperscript{51,53} Current results showed that at either short-term (4.78\% decompression vs. 3.43\% fusion) or long-term (16.00\% vs. 17.43\%) follow-up, there were no differences in reoperation rates between the two groups. However, Ghogawala et al.\textsuperscript{7} found that reoperation rate in the fusion group (14\%) was significantly lower than that in the decompression alone group (14\% in short-term follow-up (4 years). But the reoperation rate of fusion group started to increase significantly at 36 months after operation.\textsuperscript{7}

Recent RCT results showed that in short-term follow-up (2 years), reoperation frequently occurred in the decompression alone group (12.5\%), than in the fusion group (9.1\%).\textsuperscript{20} But the difference in the study by Austevoll\textsuperscript{20} was small without statistical significance. Our results are consistent with those of many previous studies.\textsuperscript{9,10,23,25,42} With regards to the included studies, the common reasons for reoperation in decompression alone group were same-segment disease (including disc herniation and recurrent stenosis), while the main reasons for reoperation in the fusion group were implant-related problems and adjacent-segment stenosis and instability.\textsuperscript{7,39} Long-term follow-up results showed that although spinal fusion accelerated the degeneration of adjacent segments, it had no effects on clinical outcomes.\textsuperscript{54} Although there was no difference at long-term follow up, reoperation rates still need to be assessed at longer follow-up for future studies. In addition, fusion was associated with longer operative time, more intraoperative blood loss, and longer hospital stay which were consistent with previous studies.\textsuperscript{7,8,20} Given the higher costs of adding fusion, decompression alone is more cost-effective than fusion.\textsuperscript{57-59}

This study had some strengths. First, this meta-analysis was performed by a professional team including a Cochrane member. Second, analyses were refined on a Patient, Intervention, Control, Outcomes, and Study design (PICOS principle). Third, our study incorporated RCTs and real-world data. However, there were some inevitable limitations in our study. First, not all of the included studies were RCTs. Real-world evidence was a useful addition. But the beliefs among surgeons may influence patient expectation and interpretation of results, so the real-world evidence may be more biased. In order to explore the bias, we conducted subgroup analyses based on study designs. Moreover, expertise of surgeons, surgical methods and inclusion/exclusion criteria in the studies included may be a potential source of heterogeneity. Then, instrumented and non-instrumented fusions were included in the fusion group. Instrumented fusions might be with higher complication rates than non-instrumented fusions which are often used to stabilize instability.\textsuperscript{51} In addition, the studies often use different definitions, different tools and different follow-up to measure outcomes like pain, complications that creates heterogeneity and biases. There is an urgent need to formulate a standard evaluation plan to measure outcomes.

This meta-analysis provides insights into evidence-based medicine currently approved by the Cochrane Collaboration.\textsuperscript{55} Our findings do not support routine applications of decompression with fusion for degenerative lumbar spondylosisisthesis.

\textbf{Contributors}

Dr X.D.Y., J.X.Q., F.L.W. and T.L. have accessed and verified the data in this study, and they take responsibility for data integrity and accuracy of analysis. F.L.W., C. P.Z. and Q.Y.G. contributed equally to this manuscript. X.D.Y., J.X.Q., F.L.W., C.P.Z., Q.Y.G. and T.L. conceived and designed the work. F.L.W., C.P.Z., Q.Y.G., M.R.D., H.R.G., K.L.Z., T.L., J.X.Q. and X.D.Y. acquired and analysed data, interpreted results. F.L.W. and C.P. Z. wrote the manuscript. All authors contributed to critical revision of the manuscript for important intellectual content. F.L.W., T.L. and C.P.Z. performed statistical analysis. X.D.Y., J.X.Q., T.L., F.L.W., C.P.Z. and Q.Y.G. contributed to administrative, technical, or material support. X.D.Y., J.X.Q., C.P.Z., F.L.W., Q.Y.G. supervised. All authors had full access to all the data in the study, and accept responsibility to submit for publication.

\textbf{Data sharing statement}

This meta-analysis of secondary analysis of raw data from published original articles. All the data used for the study are included in the manuscript and supplementary material.

\textbf{Declaration of interests}

All authors declare no competing interests.

\textbf{Acknowledgements}

This work was supported by grants from the National Natural Science Foundation of China (No. 81871818), Tangdu Hospital Seed Talent Program (Fei-Long Wei), Natural Science Basic Research Plan in Shaanxi Province of China (No.2019JM-265) and Social Talent Fund of Tangdu Hospital (No.2021SHRC034). We thank Tangdu Hospital, Fourth Military Medical University for supporting our work and Home for Researchers (www.home-for-researchers.com) for a language polishing service.

\textbf{Supplementary materials}

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.eclinm.2022.101559.
Rampersaud YR, Fisher C, Yee A, et al. Health-related quality of life following decompression compared to decompression and fusion for degenerative lumbar spondylolisthesis: a Canadian multicentre study. Can J Surg. 2014;57(4):E126–E133.

Forsth P, Michaelsson K, Sandén B. Does fusion improve the outcome after decompressive surgery for lumbar spinal stenosis? A two-years follow-up study involving 5599 patients. Bone Joint J. 2013;95-b(9):960–965.

Park JH, Hyun SJ, Roh SW, Rhim SC. A comparison of unilateral laminectomy with bilateral decompression and fusion surgery in the treatment of grade I lumbar degenerative spondylolisthesis. Acta Neurochir. 2012;154(7):1205–1212.

Kleinstück FS, Fekete TF, Mannion AF, et al. To fuse or not to fuse in lumbar degenerative spondylolisthesis: do baseline symptoms help provide the answer? Eur Spine J. 2012;21(2):268–275.

Kim S, Mortuz Hefzi S, Coyte PC, Rampersaud YR. Cost-utility of lumbar decompression with or without fusion for patients with symptomatic degenerative lumbar spondylolisthesis. Spine J. 2012;12(1):44–54.

Matsudaira K, Yamazaki T, Seichi A, et al. Spinal stenosis in grade I degenerative lumbar spondylolisthesis: a comparative study with results after following laminoplasty and laminectomy with instrumented spinal fusion. J Orthop Sci. 2005;10(1):270–276.

Ghogawala Z, Benzil EC, Amin-Hanjani S, et al. Prospective outcomes evaluation after decompression with or without instrumented fusion for lumbar stenosis and degenerative Grade I spondylolisthesis. J Neurosurg. 2004;100(3):467–272.

Yone K, Sakou T, Kawauchi Y, Yamaguchi M, Yanase M. Indication of fusion for lumbar spinal stenosis in elderly patients and its significance. Spine. 1996;21(2):242–248.

Briggs H, Milligan PR. Chip fusion of the low back following exploration of the spinal canal. J Bone Joint Surg Am. 1954;36-A:125–130.

Weiss AJ EA, Andrews RM. Characteristics of Operating Room Procedures in US Hospitals, 2011: Statistical Brief #170. Healthcare Cost and Utilization Project (HCUP) Statistical Briefs. Rockville, MD: Agency for Healthcare Research and Quality (US); 2014.

Restrick DK, Watters WC, yd, Sharan A, et al. Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 9: lumbar fusion for stenosis with spondylolisthesis. J Neurosurg. 2012;21(1):54–61.

Johnsson KE, Redlund-Johnell I, Udén A, Willner S. Preoperative and postoperative instability in lumbar spinal stenosis. Spine. 2001;26(14):591–593.

Mannion AF, Leivseth G, Brox JJ, Fritzeell P, Hägg O, Fairbank JC. ISSLS Prize winner: Long-term follow-up suggests spinal fusion is associated with increased adjacent segment disc degeneration but without influence on clinical outcome: results of a combined follow-up from 4 randomized controlled trials. Spine. 2014;39(17):1373–1383.

Packer M. Are meta-analyses a form of medical fake news? Thoughts about how they should contribute to medical science and practice. Circulation. 2017;136(22):2097–2099.