Artificial Total Disc Replacement Versus Fusion for Lumbar Degenerative Disc Disease: An Update Systematic Review and Meta-Analysis

Yu-Zhe LI, Piao SUN, Dong CHEN, Li TANG, Chun-Hui CHEN, Ai-Min WU

The Second Affiliated Hospital and Yuying Children’s Hospital of Wenzhou Medical University, The Second School of Medicine Wenzhou Medical University, Zhejiang Spine Surgery Center, Department of Spine Surgery, Wenzhou, China

Corresponding author: Ai-Min WU  aiminwu@wmu.edu.cn

ABSTRACT

AIM: To conduct an updated systematic review and meta-analysis to compare the efficacy and safety between total disc replacement (TDR) and fusion surgery for lumbar degenerative disc disease (LDDD).

MATERIAL and METHODS: We comprehensively searched meta-analyses comparing TDR with fusion through the PubMed, Embase, and Cochrane Library databases. Only randomized controlled trials (RCTs) were selected and collected. The end of the retrieval time was June 2017. Two authors independently extracted the data from the studies after assessing their quality. The statistical software STATA version 12.0 was used to analyze the data.

RESULTS: A total of seven RCTs (1706 patients) were included in our analysis. The patients in the TDR group had significantly improved. A greater percentage of these patients were satisfied with the surgery concerning Oswestry disability index, visual analog scale score, and complication rate. In addition, the clinical success in the TDR group was greater than that in the fusion group. Meanwhile, the TDR group had shorter operative time and hospital stay. However, there was no clinical significance regarding blood loss, work status, and reoperation rate between the two groups.

CONCLUSION: Our current updated meta-analysis suggests that TDR could be an alternative treatment for LDDD, since it yielded better clinical success and patient satisfaction, shorter hospital stay and operative time, less pain, and lower complication rates than lumbar fusion.

KEYWORDS: Lumbar degenerative disc disease, Lumbar total disc replacement, Lumbar fusion, Systematic review, Meta-analysis

INTRODUCTION

Lumbar degenerative disc disease (LDDD) includes lumbar stenosis, instability, and disc herniation. The conservative treatment for LDDD includes administration of non-steroidal anti-inflammatory drugs and physical therapy. If this treatment does not yield good outcomes, surgical interventions will be required (5). Lumbar fusion has been developed for several decades and regarded as the gold standard treatment for LDDD (2,8,20). On radiography, the fusion rate has reached 96.7% and clinical satisfaction rate, 65%-95% (9,32). However, this treatment has disadvantages, such as stiffness of the surgical levels, and may increase the rate of adjacent level diseases (19,21,34).

In recent years, artificial total disc replacement (TDR), also named arthroplasty, in which a degenerated lumbar disc is replaced with a moveable prosthesis to treat LDDD, has been developed as an alternative technique and received appreciable attentions. This technique can mimic the motion range of the natural disc and retain the original biomechanical function of the lumbar spine (1). Patients’ normal intervertebral
segment motions could be restored because the adjacent levels circumvent the nonphysiologic loading after TDR (18,38,40). Previous meta-analyses have drawn different conclusions regarding whether TDR has a significant superiority for treating LDDD when compared with lumbar fusion (33,41). Wu et al. reported the results of 837 patients from five trials in 2010 (41), which indicated that there was no significant difference between TDR and lumbar fusion for reducing pain. Rao and Cao presented the results of 1584 patients from seven trials in 2014 (33); they found that the safety and efficacy, including that in pain reduction, of TDR were comparable with those of lumbar fusion at 2 years of follow-up.

Recently, some randomized controlled trials (RCTs), including remarkable RCTs with a long-term follow-up (5 years), have compared between TDR and lumbar fusion for treating LDDD. We systematically searched the PubMed, Embase, and Cochrane Library databases to obtain the current and best available evidence and refined the indicators for RCTs, which can reflect on the safety during the perioperative and postoperative periods, feedback from patients, and clinical results. Therefore, we performed this present updated systematic review and meta-analysis in a more comprehensive manner.

■ MATERIALS and METHOD

This was a systematic review and meta-analysis, without identified information of the primary patients; thus, ethical approval was not necessary. Further, our present systematic review and meta-analysis was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (Checklist S1).

Literature Search Strategy

We systematically searched the PubMed, Embase, and Cochrane Library databases for articles published until June 2017, with the limitation of English language. The search keywords included as follows: lumbar disc replacement, lumbar disc arthroplasty, lumbar arthroplasty, and lumbar fusion. All RCTs that compared TDR with lumbar fusion for LDDD were collected for the analysis.

Inclusion Criteria

The inclusion criteria were as follows: 1) RCTs comparing lumbar disc replacement with lumbar fusion for LDDD; 2) number of cases studied in the literature, ≥15; 3) follow-up period, ≥2 years; 4) no history of surgical treatment other than the treatment studied; and 5) symptomatic degenerative lumbar diseases.

Data Extraction

The two authors independently extracted the following data from the included studies: first author, publication year, publication journal, total number of cases, age and sex of the patients, type of prosthesis, and follow-up period. The outcomes of the clinical trials extracted included the Oswestry disability index (ODI), visual analog scale (VAS) score, rate of complications, operative time, blood loss, hospital stay, and rate of reoperation.

Quality assessment:

The Cochrane Handbook for Systematic Reviews of Interventions was used for the methodological quality assessment in the study. We assessed the selected study using the following characteristics (12).

1. Was the allocation sequence adequately generated?
2. Was allocation adequately concealed?
3. Was blinding used in the participant, personnel, and outcome assessment?
4. Were incomplete outcome data adequately addressed?
5. Were reports of the study free of suggestion of selective outcome reporting?
6. Was the study apparently free of other problems that could put it at a risk of bias?

If the answer for a judgment criterion was “YES,” we considered it to indicate a low risk of bias; “NO,” high risk of bias; “UNCLEAR,” unclear risk of bias. The quality of the study was divided into three levels: A, all the risks of bias were low; B, the criteria in A were partially met; and C, no criteria in A were met.

Heterogeneity

The clinic homogeneity was interrelated to several characteristics of the participants, including sex, age, clinical manifestation, functional status at baseline, and pain. The surgical technique homogeneity included the type of artificial lumbar disc and fusion method, measurement method, follow-up period, and exclusion criteria. The chi-squared test was employed to identify the heterogeneity of the outcomes, which can describe that the proportion of variation in the outcomes was caused by heterogeneity rather than random change. The I² ranged from 0% to 100%, with 0% indicating no heterogeneity. When the I² was <50%, the heterogeneity was low, and the fixed effect model was used. When the I² was >50%, the heterogeneity was high, and the random effect model was used.

■ RESULTS

The main characteristics of the identified relevant studies are summarized in Table I. From the data bases searched, 622 references were obtained. The two authors scanned both the titles and abstracts, and 598 irrelevant references were excluded. In the remaining 24 potentially relevant references, 18 were further excluded, since they did not meet the inclusion criteria. The flow of study selection is described in Figure 1. Finally, seven RCTs that compared the 2- or 5-year follow-up results between TDR and lumbar fusion were included. These seven relevant RCTs included 1706 patients with degenerative lumbar diseases: 1150 patients in the TDR group and 556 patients in the fusion group (6,11,14,15,36,37,45).

Risk of Bias

The outcomes of the included studies are presented in Table II. Almost all studies included achieved a high quality
according to the Cochrane Handbook for Systematic Reviews; however, most of them did not employ the blinding method (6,14,36,45). This prominent limitation may cause a certain degree of detection bias. All of the participants in the seven studies underwent follow-up for at least 2 years; three of the studies had 5 years of follow-up. A follow-up rate of >89% was obtained in all studies, except in that by Sasso et al. (36). However, we still included data from the study by Sasso et al., which were irrelevant to the follow-up studies, to obtain a more comprehensive perspective.

Table I: Characteristics of Seven Randomized Controlled Trials (RCTs) Included in the Present Study

<table>
<thead>
<tr>
<th>Author (year)</th>
<th>Number of patients (TDR / F)</th>
<th>Age (Mean) (years) (Mean ± SD) (years) (TDR / F)</th>
<th>Gender Male (TDR / F)</th>
<th>Follow-up (years)</th>
<th>Type of TDR</th>
<th>Surgical approach for lumbar fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blumenthal et al. (2005) (6)</td>
<td>205/99</td>
<td>39.6 / 39.6</td>
<td>113/44</td>
<td>2 years</td>
<td>Charite’</td>
<td>anterior</td>
</tr>
<tr>
<td>Sköld et al. (2013) (37)</td>
<td>80/72</td>
<td>40.2 ± 8.1 / 38.5 ± 7.8</td>
<td>32/30</td>
<td>5 years</td>
<td>Maverick</td>
<td>posterior</td>
</tr>
<tr>
<td>Zigler et al. (2012) (45)</td>
<td>161/75</td>
<td>38.7 / 40.4</td>
<td>82/34</td>
<td>5 years</td>
<td>Prodisc-l</td>
<td>anterior</td>
</tr>
<tr>
<td>Geisler et al. (2009) (14)</td>
<td>90/43</td>
<td>39.96 / 38.7</td>
<td>59/27</td>
<td>5 years</td>
<td>Charite’</td>
<td>anterior</td>
</tr>
<tr>
<td>Sasso et al. (2008) (36)</td>
<td>44/23</td>
<td>36 / 41</td>
<td>23/10</td>
<td>2 years</td>
<td>Flexilore</td>
<td>anterior/posterior</td>
</tr>
<tr>
<td>Gonet et al. (2011) (15)</td>
<td>405/172</td>
<td>39.9 / 40.2</td>
<td>205/86</td>
<td>2 years</td>
<td>Maverick</td>
<td>posterior</td>
</tr>
<tr>
<td>Delamarter et al. (2011) (11)</td>
<td>165/72</td>
<td>41.8 ± 7.73 / 41.8 ± 7.81</td>
<td>95/39</td>
<td>2 years</td>
<td>Prodisc-l</td>
<td>anterior</td>
</tr>
</tbody>
</table>

TDR: Total disc replacement, F: Fusion, SD: Standard deviation.

Figure 1: Selection of identified studies for the systematic review and meta-analysis.
**Operative Time**
The operative time was reported in all seven studies (6,11,14,15,36,37,45); significant heterogeneity existed across seven trials ($I^2=98.0\%$, $p=0.000$), and the random effect model was employed. Compared with the patients treated with fusion, the patients treated with TDR had a significantly decreased operative time (SMD=$-1.16$; 95%CI=$-1.98$ to $-0.35$; $p=0.005$) (Figure 2).

**Hospital Stay**
The hospital stay was reported in all seven trials (Figure 2) (6,11,14,15,36,37,45). Significant heterogeneity also existed across all trials ($I^2=96.2\%$, $p=0.000$), and the random effect model was used; the TDR-treated patients had a significantly decreased hospital stay compared with the fusion-treated patients (SMD=$-0.95$; 95%CI=$-1.55$ to $-0.35$; $p=0.002$).

**Blood Loss**
All seven trials reported blood loss (Figure 2) (6,11,14,15,36,37,45). There was also significant heterogeneity across them ($I^2=95.9\%$, $p=0.000$), and the random effect model was employed. There was no significant difference in blood loss between the patients treated with TDR and fusion (SMD=$-0.48$; 95%CI=$-1.03$ to $-0.07$; $p=0.09$).

**VAS Score**
Six trials reported the VAS score (6,11,14,15,36,37,45); no significant heterogeneity existed among them ($I^2=0.0\%$, $p=0.630$). Compared with the patients treated with lumbar fusion, the patients treated with artificial TDR showed a significantly decreased VAS score for their leg or back pain (SMD=$-0.18$; 95%CI=$-0.29$ to $-0.08$; $p=0.001$) (Figure 3).

**ODI**
Six trials reported the ODI (6,11,14,15,36,37,45); a low heterogeneity existed among these studies ($I^2=36.7\%$, $p=0.136$). The patients’ functional ability as indicated by the ODI in the TDR group (SMD=$-0.20$; 95%CI=$-0.34$ to $-0.05$; $p=0.007$) was significantly better than that in the fusion group (Figure 3).

**Work Status**
The work status refers to the percentage of patients who participate in full-time or part-time jobs (Figure 4). It was reported in four trials (11,14,15,45), with 1183 patients (821 patients in the TDR group and 362 patients in the fusion group). A low heterogeneity existed across the four studies ($I^2=38.3\%$, $p=0.182$), and the fixed effect model was used. There was no significant difference in the work status between the patients treated with TDR and fusion (RR=$1.03$; 95% CI=$0.95$ to $1.12$; $p=0.46$).

**Clinical Success**
Five trials reported the clinical success (11,14,15,37,45). The rate of clinical success in the TDR group was 79.7% (810/1016), and that in the fusion group was 72.3% (347/480). There was significant heterogeneity across the five trials ($I^2=52.4\%$, $p=0.078$), and the random effect model was employed (Figure 4). The clinical success in the TDR group was significantly greater than that in the fusion group (RR=$1.10$; 95% CI=$1.03$ to $1.17$; $p=0.003$).

**Satisfaction with the Surgery**
Six trials reported the responses of the patients regarding whether they were satisfied with the surgery (6,11,14,15,37,45). The percentage of the patients who answered “yes” in the TDR group was 75.5% (835/1106), and that in the fusion group was 64.2% (335/522). There was low heterogeneity across the six trials ($I^2=44.2\%$, $p=0.111$), and the fixed effect model was used (Figure 4). The satisfaction rate was higher in the TDR group than in the fusion group (RR=$1.18$; 95% CI=$1.10$ to $1.27$; $p=0.000$).

**Complications**
The major complications reported included dural tear, iliac

---

**Table II: Risk of Bias Assessment of All Included Studies**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
</tr>
<tr>
<td>Allocation concealment</td>
<td>Low risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
</tr>
<tr>
<td>Blinding of participants and personnel</td>
<td>High risk</td>
<td>High risk</td>
<td>High risk</td>
<td>High risk</td>
<td>High risk</td>
<td>High risk</td>
<td>High risk</td>
</tr>
<tr>
<td>Blinding of outcome data addressed</td>
<td>High risk</td>
<td>High risk</td>
<td>High risk</td>
<td>High risk</td>
<td>High risk</td>
<td>High risk</td>
<td>High risk</td>
</tr>
<tr>
<td>Incomplete outcome data addressed</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>High risk</td>
<td>Low risk</td>
<td>Low risk</td>
</tr>
<tr>
<td>Selective reporting</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
<td>Low risk</td>
</tr>
<tr>
<td>Free of other bias</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
<td>Unclear risk</td>
</tr>
</tbody>
</table>
Figure 2: Forest plot showing the meta-analysis of the operative time, hospital stay, and blood loss in the TDR and fusion groups. The TDR-treated patients had significantly decreased operative time and hospital stay compared with the fusion-treated patients. TDR, total disc replacement.

Figure 3: Forest plot showing the meta-analysis of the visual analog scale score and Oswestry disability index in the TDR and fusion groups. The results showed that the TDR-treated patients had a significant improvement in both indicators. TDR, total disc replacement.
The complications were reported in five trials (6,14,36,37,45). The complication rate was 15.3% (100/655) in the TDR group and 26.8% (86/321) in the fusion group. There was a low heterogeneity across the five trials ($I^2=16.6\%$, $p=0.309$), and the fixed effect model was employed (Figure 5). There were fewer complications in the TDR group than in the fusion group (RR=0.59; 95% CI=0.47 to 0.75; $p=0.000$).

**Reoperation Rate**

The reoperation rate, referring to the rate of secondary surgical procedures performed to correct, remove, or reoperate the implantor fusion fixation, was reported in six trials (6,11,15,36,37,45) with 1505 patients (1028 TDR-treated patients and 477 fusion-treated patients). The reoperation rate in the TDR group was 6.3% (65/1028), and that in the fusion group was 8.4% (40/477). There was no heterogeneity across the six trials ($I^2=1.1\%$, $p=0.409$), and the fixed effect model was used (Figure 6). There was no significant difference in the reoperation rate between the TDR group and the fusion group (RR=0.76; 95% CI=0.52 to 1.11; $p=0.152$).

**DISCUSSION**

Lumbar fusion remains the established standard technique in the treatment for LDDD (3,10,28,29). The disadvantages of the fusion technique include adjacent segmental degeneration and stiffness of the surgical levels and altered original biomechanics of the spine (17,22,42,44). Reoperation is also generally required because of adjacent segmental degeneration (30). As a new non-fusion treatment, TDR has been extensively used by surgeons as an alternative option for treating LDDD (39).

TDR was designed to maintain and restore the spinal segment motion; it may have the potential advantage of preventing adjacent level degeneration (13,26). However, whether TDR is significantly superior to the golden standard lumbar fusion remains unclear. The purpose of this study is to compare the efficacy and safety between TDR and lumbar fusion in the treatment of LDDD. A previous meta-analysis has been performed to accomplish this purpose but was not able to draw a convincing conclusion. Our study aimed to deduce a more virtuous conclusion by adopting long-term follow-up studies, which could be considered as the current best available evidence.

The seven trials contained three perioperative period indicators, including operative time, blood loss, and hospital stay. The operative time and hospital stay in the TDR group significantly decreased when compared with those in the fusion group; blood loss showed no significant difference between the two groups; only Gonet et al. reported that the operative time was longer, and the blood loss was greater in
the TDR group (15). The conception that TDR was a difficult and new procedure for surgeons could explain the longer operative time and greater blood loss in their study.

This meta-analysis showed that TDR yielded better clinical outcomes, particularly the ODI and VAS score, than did lumbar fusion. We obtained results similar to those of a previous meta-analysis that compared the VAS score and ODI between TDR and fusion; however, the difference is our study included three 5-year follow-up trials (14,37,45). Geisler et al. reported contrary results on the VAS score and ODI; they found that the fusion group had better clinical outcomes (14). This may be because the trial was performed at a time when TDR was not a mature treatment; the article was published in 2008; and the time of the treatments for the participants was between May 2000 and April 2002.

Six trials reported the responses of the patients regarding whether they were satisfied with the surgery (6,11,14,15,37,45). A low heterogeneity was found. It seemed that TDR was accepted and welcomed by the patients.

Figure 5: Forest plot showing the meta-analysis of the complications in the TDR and fusion groups. There were fewer complications in the TDR group than in the fusion group. TDR, total disc replacement.

Figure 6: Reoperation rate in the total disc replacement and fusion groups. There was no significant difference between the two groups.
The reoperation rate and complications are used to assess the safety of surgeries. Herein, there were significantly fewer complications in the TDR group than in the fusion group in five included trials (6,11,36,37,45); this finding was different from those of a previous meta-analysis. This implies that TDR yields a better prognosis than does fusion on a long-term basis. However, the reoperation rate showed no significant difference in six included trials (6,11,15,36,37,45), which is similar to that of a previous meta-analysis.

Although the TDR and fusion groups had a similar rate of complications, one major concern is that the TDR group may have some complications distinct to those in the lumbar fusion group. TDR was reported to have yielded acute S1 radiculopathy potential device-related complications and retroperitoneal lymphocele (16,27). TDR was performed via anterior approaches to the lumbar spine and required retraction of the great vessels and visceral organs, which could cause ureteral injury (24,43); however, Parks et al. showed that TDR does not increase the rate of hydronephrosis (31). The rate of retrograde ejaculation after TDR was lower than that after lumbar fusion (4,23). In addition, Malham and Parker reported that the regional risks of vascular and visceral injuries after traditional TDR performed via the anterior retroperitoneal approach can be avoided by adopting the direct lateral retroperitoneal transpsoas approach (25). As TDR has a lower reoperation rate but yields concerns regarding the structural integrity of the anterior spinal column, which increase the difficulty of the reoperation, surgeons should be attentive and careful while planning or performing reoperations (7).

The statistical results of our study matched those of a previous meta-analysis in most indicators selected.

“The complication rate and operative time were better in the TDR group in this study than in the previous meta-analysis by Rao and Cao in 2014 (33). As such, TDR may show a significant superiority for chronic complications, and most surgeons have become familiar with the procedure of this new technology. However, a publication bias could result in the overestimation of the difference between the two treatments, and more remarkable RCTs with long-term follow-ups are needed to verify the results.

Although the current evidence showed the superiority of TDR to lumbar fusion, its use remains limited. The reasons for this may include as follows: 1) most surgeons are familiar with the posterior lumbar fusion; it would be difficult to change their preference; 2) the anterior TDR approach involves many complex vessels and nerve tissues; soft tissue injuries will induce catastrophic results, even death (24,43). Whether TDR can decrease the rate of adjacent level diseases remains unclear; 4) the management of the complications is difficult; and 5) the cost of TDR is higher than that of fusion; the refusal of some insurance companies to reimburse TDR is also one of the reasons why its clinical use is limited (35).

Strengths

In this meta-analysis, all of the included studies were RCTs, which can avoid any selection bias. The methodological qualities of six studies were satisfactory (the quality of all studies was above level C). All included studies conducted follow-ups of >2 years; specifically, three of them had a follow-up duration of >5 years. The total number of randomized participants was 1706, which is considered very large in a spinal implant study. Therefore, the results could be considered credible. In addition, we selected our indicators comprehensively to contrast the safety during the perioperative and postoperative periods, feedback from patients, and clinical results, which were not presented explicitly in previous meta-analyses.

Limitations

There were several limitations in the present study. First, the methodological limitation appeared to be inadequacy regarding the outcome assess or blinding to the intervention, and the type of implant could be recognized obviously on the postoperative radiographic films. Second, the data on the operative time, hospital stay, and blood loss had significant heterogeneity; thus, the random effect model was employed; these results should then be interpreted with caution.

CONCLUSIONS

Our current updated meta-analysis suggests that TDR could be an alternative treatment for LDDD, since it yielded better clinical success and patient satisfaction, shorter hospital stay and operative time, less pain, and lower complication rates than did lumbar fusion.

ACKNOWLEDGEMENTS

This work was supported by Wenzhou leading talent innovative project (RX2016004), Zhejiang Provincial Medical Technology Foundation of China (2018KY129), National Natural Science Foundation of China (81501933) and Wenzhou Municipal Science and Technology Bureau (Y20170389). The funders had no role in the design, execution, or writing of the study.

REFERENCES


