Comparison of Total Disc Replacement with Lumbar Fusion: A Meta-Analysis of Randomized Controlled Trials

Hongfei Nie, Guo Chen, Xiandi Wang and Jiancheng Zeng

Abstract

A meta-analysis was performed to evaluate whether a beneficial clinical effect of the Total Disc Replacement (TDR) over lumbar fusion for the treatment of patients with Degenerative Disc Disease (DDD). An electronic search of PubMed, Cochrane Central Register of Controlled Trials, and EMBASE from their inception to 2012 was completed, and we assessed risk bias and retrieved relevant data, and meta-analysis was performed, if appropriate. Oswestry Disability Index (ODI), Visual Analog Score (VAS), patient satisfaction or VAS patient satisfaction, narcotic use, overall success rate, reoperation rate, work status, “surgery again?”, complications and radiographic outcomes were evaluated. Six RCTs were included in this meta-analysis. At 2 years, TDR was demonstrated to be more beneficial for patients compared to lumbar fusion in the following outcomes, including ODI scores (MD:-4.87, 95% CI: -7.77 to -1.97, p=0.001), patient satisfaction (OR:1.91, 95% CI: 1.27 to 2.86, p=0.002) and VAS patient satisfaction (MD:9.10, 95% CI: 3.20 to 14.99, p=0.002), the percentage of using narcotics (OR=0.54, 95%CI: 0.31 to 0.96, p=0.03), overall success rate (OR:1.68, 95% CI: 1.26 to 2.25, p=0.005), the rate of patients to chose the same surgical treatment again (OR:2.38, 95% CI: 1.72 to 3.28, p < 0.001), and complications (OR=0.50, 95%CI: 0.29 to 0.84, p=0.008). Other outcomes, including re-operation rate (OR:0.62, 95% CI: 0.36 to 1.06, p=0.08) and work status (OR=1.05, 95% CI: 0.75 to 1.47, p=0.80), were demonstrated to be no differences between the two groups. In a long-term of follow-up (2 years), TDR shows a significant superiority for the treatment of lumbar DDD compared with fusion.

Key Words: *Total Disc Replacement (TDR). Lumbar fusion. Degenerative disc disease. Re-operation rate.*

INTRODUCTION

Degenerative Disc Disease (DDD) can lead to disc dehydration, annular tears, and/or loss of disc height or collapse, and can result in abnormal motion of the segment and biomechanical instability of the spine.[1] DDD has been the leading cause for chronic low back pain and dysfunction in the society.[2] In patients suffering from chronic low back pain caused by DDD, previous studies have shown that surgical intervention has benefits over conservative treatment for debilitating the low back pain.[3,4] The rationale for surgical treatment in DDD has long been based on the idea that limiting motion of a pain-producing segment will limit the pain generated by that segment.[5]

According to this, lumbar fusion is considered as an effective treatment for patients with DDD to eliminate abnormal motion and eliminate instability at the symptomatic degenerated levels, and, therefore, reduce the low back pain.[6] Although fusion surgery yields better results in decreasing pain and disability compared to the conservative treatment, it also has detrimental effects on the normal physiological and biomechanical function of the spine.[5] As decreased mobility of the painful degenerative segment could lead to increased stress on the neighbouring segment, fusion is often associated with future degeneration at the adjacent levels.[7,8] As a result, the need for non-fusion techniques is on the rise.

Total lumbar Disc Replacement (TDR) has been shown to be a promising alternative in treatment of low back pain caused by DDD, and may reduce the bio-mechanical changes associated with fusion through restoring the disc height and preserving segmental motion after removing the source of nerve root or spinal cord compression.[9] The mechanism of pain relief is based on a combination of complete excision of the painful disc and restoration of segmental load transfer, sagittal balance and motion.[10,11] Besides, a secondary intention of this technique is the preservation of normal motion at the adjacent lumbar levels, hoping that this will reduce later degeneration of the adjacent lumbar segments.[12]

Previous Randomized Clinical Trials (RCTs) of TDR at one or two level demonstrated results that were equivalent or superior to those of lumbar fusion at 2 years of follow-up.[6,9,13-16] However, the long-term clinical outcomes of surgically treated DDD with TDR or lumbar fusion have not been entirely studied. Therefore, the objective of this study is to systematically search relevant RCTs and to comprehensively compare the

Department of Orthopedics, West China Hospital, Sichuan University, Chengdu, Sichuan, China.

Correspondence: Dr. Jiancheng Zeng, Department of Orthopedics, West China Hospital, Sichuan University, 17#, Third Segment, Renmin South Road, Chengdu, Sichuan 610041, China.
E-mail: savagenhf@gmail.com

Received: September 19, 2012; Accepted: July 29, 2014.

long-term clinical outcomes of TDR with lumbar fusion for the treatment of patients with DDD.

Search strategy: All randomized controlled clinical trials (RCTs) comparing the TDR to fusion for the treatment of lumbar degenerative disc disease were identified in this study. Electronic databases including PubMed (1966 to September 2011), EMBASE (1984 to September 2011), Cochrane Controlled Trials Register (Central; 3rd Quarter, 2011) were searched. The search strategy consisted of a combination of keywords concerning the technical procedure (total disc replacement, prosthesis, implantation, discectomy, arthroplasty) and keywords regarding the anatomical features and pathology (lumbar vertebrae). These keywords were used as MeSH headings and free text words. In addition, a search was performed using the specific names of the prostheses. We identified all relevant RCTs, searched reference lists of review articles, and included studies to identify other potentially eligible studies.

Selection of studies: Two review authors independently examined all titles and abstracts that met the search terms and reviewed full publications, when necessary. The reference section of all primary studies was inspected for additional references, and only those reporting the results of a randomized controlled trial were included in this analysis. The search was limited to studies published in English, and only trials with 2-year follow-up results reported were included in this meta-analysis. This review was conducted under the suggested Quorum guideline standards.[17] If studies did not report the actual number or the standard deviation but rather presented the data only in graph format, the authors were contacted. Most authors responded but were not able to provide additional clarification because of personal circumstances, or because the data presented were preliminary and not available for scientific research, and thus these studies were excluded.

Data extraction: Two review authors independently extracted relevant data from the included studies regarding the design, age, gender, types of prosthesis and length of follow-up. For each trial, the clinical outcomes were collected in terms of the improvement of movement and functioning measured by a disability scale (Oswestry Disability Index (ODI)); the improvement in pain measured by a validated pain scale (Visual Analog Score (VAS)); patient satisfaction or VAS patient satisfaction; the rate of narcotic usage for pain; overall success rate; the reoperation rate for secondary surgery; patients' work status; the rate of patients would have the same surgical treatment again; complications; Range of Motion (ROM) and disc height.

Heterogeneity: To establish inconsistency in the study results, the test for heterogeneity (Cochrane Q) was performed. However, because the test is susceptible to the number of trials included in the meta-analysis, I^2 was directly calculated from the Q statistic, which describes the percentage of variation across the studies that is due to heterogeneity rather than change.

Assessment of risk bias: Two independent investigators evaluated the risk bias of the included studies. Briefly, as the risk of overestimation of intervention effects in RCTs with inadequate methodology [18 - 20], we assessed the influence of risk bias using the following components; randomization and generation of the allocation sequence; allocation concealment; blinding; and description of the follow-up. The details of each methodological item are shown in Table II. Due to the nature of surgical treatment, the domain of blinding could not be easily performed, and, therefore, the trials with an adequate method of allocation sequence and allocation concealment as well as clearly description of the follow-up were considered to be with high quality.

A meta-analysis was conducted using the software Revman 5.1 (provided by the Cochrane Collaboration, Oxford, UK) for an outcome where data are available from more than one study. The analyses included all patients irrespective of compliance or follow-up following the "intention-to-treat" principle and using the last reported observed response. We presented dichotomous variables as Odds Ratios (OR) with 95% Confidence Interval (CI) and continuous outcomes as Mean Differences (MD) with 95% CI. The fixed effects model and the random effects model were used, with the

Table I: Main characteristics of included studies.

Trials	Number of patients	Age (year)	Male (%)	Type of disc	Follow-up (years)	Related outcomes
Blumenthal	205/99	39.6/39.6	55.1/44.4	Charité disc	2	ODI, VAS, PS, WS, NU, OS, reoperation rate, complications, SA, ROM, Disc height
Delamarter	56/22	39.7/44.2	57.0/45.0	ProDisc	2	ODI, VAS
Zigler	161/75	38.7/40.4	50.9/45.3	ProDisc	2	ODI, VAS, SF-36, WS, NU, OS, reoperation rate, complications, SA, ROM
Sasso	44/23	36.0/41.0	52.3/43.5	FlexiCore Disc	2	ODI, VAS, complications, ROM
Berg	80/72	40.2/38.5	40.0/42.0	ProDisc Charité disc Maverick	2	ODI, VAS, SF-36, EQ5D, PS, WS, OS, reoperation rate, complications, ROM, disc height
Delamarter	165/72	41.8/41.8	57.6/54.2	ProDisc	2	ODI, VAS, SF-36, WS, NU, OS, reoperation rate, complications, SA, ROM

ODI: Oswestry disability index, VAS: Visual analog score, PS: Patient satisfaction or VAS patient satisfaction, WS: Work status, NU: Narcotic usage, OS: Overall success rate, SA: Surgery again, ROM: Range of motion.

Table II: Risk bias of the included studies.

Trials (year)	Randomization	Patient blinding	Examiner blinding	Withdrawals and dropouts	Allocation concealment
Blumenthal (2005)	Yes / adequate	No use	No use	Clear report	Adequate
Delamarter (2005)	Yes / adequate	No use	No use	Clear report	Unclear
Zigler (2007)	Yes / adequate	No use	No use	Clear report	Unclear
Sasso (2008)	Yes / unclear	No use	No use	Clear report	Unclear
Berg (2009)	Yes / unclear	No use	No use	Clear report	Adequate
Delamarter (2011)	Yes / adequate	No use	No use	Clear report	Unclear

The details of each methodological item

Randomization:

- Adequate, if the allocation sequence was generated by a computer or random number table. Drawing of lots, tossing of a coin, shuffling of cards, or throwing dice will be considered as adequate if a person who was not otherwise involved in the recruitment of participants performed the procedure.
- Unclear, if the trial was described as randomised, but the method used for the allocation sequence generation was not described.
- Inadequate, if a system involving dates, names, or admittance numbers were used for the allocation of patients.

Blinding:

- Adequate, if the trial was described as double blind and the method of blinding involved identical placebo or active drugs.
- Unclear, if the trial was described as double blind, but the method of blinding was not described.
- Not performed, if the trial was not double blind.

Withdrawals and dropouts:

- Adequate, if the numbers and reasons for dropouts and withdrawals in all intervention groups were described or if it was specified that there were no dropouts or withdrawals.
- Unclear, if the report gave the impression that there had been no dropouts or withdrawals, but this was not specifically stated.
- Inadequate, if the number or reasons for dropouts and withdrawals were not described.

Allocation concealment:

- Adequate, if the allocation of patients involved a central independent unit, on-site locked computer, identically appearing numbered drug bottles or containers prepared by an independent pharmacist or investigator, or sealed envelopes.
- Unclear, if the trial was described as randomised, but the method used to conceal the allocation was not described.
- Inadequate, if the allocation sequence was known to the investigators who assigned participants or if the study was quasi-randomised.

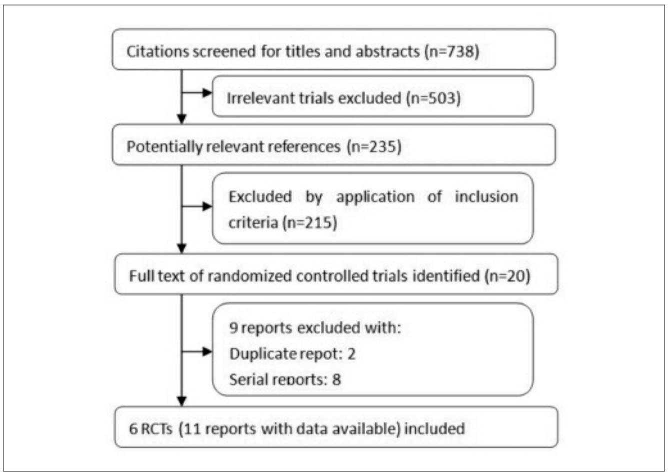

Figure 1: Flow of study identification, inclusion, exclusion.

significant level set at p=0.05. In addition, we planned to use funnel plot asymmetry to assess the existence of publication bias and other biases.[21]

Figure 1 shows the details of study identification, inclusion, and exclusion. The search on PubMed, EMBASE and the Cochrane Library under the defined terms yielded 738 articles. By screening the titles and abstracts, 503 references were excluded due to the irrelevance to this topic. In 235 potentially relevant references, 215 references were excluded and the remaining 20 reports were taken for a comprehensive evaluation. Finally, 11 reports from 6 RCTs were included in this meta-analysis.[6,9,13-16] The main characteristics of included studies was shown in Table I. Six included RCTs enrolled 1074 patients with one or two level of lumbar disc disease. Seven hundred eleven patients were randomizedly assigned into TDR group, while the other 363 patients assigned into fusion group. The studied lumbar disc prosthesis included ProDisc-L system (Synthes Spine, West Chester, PA), Charité disc system (DePuy Spine, Raynham, MA), FlexiCore disc system (Stryker Spine, Allendale, NJ) and Maverick System (Medtronic, Memphis, TE).

Risk of bias in these trials: The authors were unable to perform the funnel plot analysis as stated in the protocol, as both visual examination and statistical analysis of funnel plots have limited power to detect bias if the number of trials is small. Most of the included trials reported the power calculations to assess the sample size, except the Delamarter's study[9] and the Sasso's study.[14] Generation of the allocation sequence was considered adequate in four trials.[6,9,13,16] while in the other two trials it was either not described or unclear.[14,15] Sealed-envelop technique for allocation concealment was applied in two trials.[6,15] Blinding was not performed in all trials, while the follow-up was considered adequate in all included trials. Quality assessment reveals that all included studies except the Blumenthal's study,[6] with two or more unclear or inadequate quality components, were, therefore, regarded as high-bias risk trials (Table II).

Table III: Meta-analysis results of total disc replacement versus lumbar fusion.

Outcomes	Trials (N)	Pooled estimates			Heterogeneity		
A: including the study with stand-alone cage interbody fusion		MD/OR 95% CI	Z	p-value	x^2	p	I^2,%
ODI scores	5 trials (n=1007)	-4.87 [-7.77, -1.97]	3.29	0.001	1.59	0.81	0
VAS pain scores	5 trials (n=1007)	-5.13 [-9.02, -1.25]	2.59	0.01	0.50	0.97	0
Patient satisfaction	2 trials (n=456)	1.91 [1.27, 2.86]	3.13	0.002	2.39	0.12	58
VAS patient satisfaction	2 trials (n=473)	9.10 [3.20, 14.99]	3.03	0.002	0.01	0.92	0
Narcotic use	3 trials (n=743)	0.54 [0.31, 0.96]	2.12	0.03	5.19	0.07	61
Overall success rate	4 trials (n=884)	1.68 [1.26, 2.25]	3.50	0.0005	2.21	0.53	0
Reoperation rate	4 trials (n=929)	0.62 [0.36, 1.06]	1.74	0.08	2.43	0.49	0
Work status	4 trials (n=892)	1.05 [0.75, 1.47]	0.26	0.80	4.13	0.25	27
Surgery again	4 trials (n=818)	2.53 [1.57, 4.06]	3.83	0.0001	5.60	0.13	46
Complications	4 trials (n=692)	0.50 [0.29, 0.84]	2.63	0.008	2.62	0.45	0
B: excluding the study with stand-alone cage interbody fusion		MD/OR 95% CI	Z	p-value	x^2	p	I^2,%
ODI scores	4 trials (n=703)	-5.10 [-8.54, -1.67]	2.91	0.004	1.53	0.68	0
VAS pain scores	4 trials (n=703)	-4.90 [-9.50, -0.30]	2.09	0.04	0.47	0.93	0
Patient satisfaction	1 trial (n=152)	1.24 [0.62, 2.47]	0.61	0.54	NA	NA	NA
VAS patient satisfaction	2 trials (n=473)	9.10 [3.20, 14.99]	3.03	0.002	0.01	0.92	0
Narcotic use	2 trials (n=439)	0.61 [0.26, 1.42]	1.15	0.25	4.20	0.04	76
Overall success rate	3 trials (n=580)	1.91 [1.33, 2.75]	3.49	0.0005	0.86	0.65	0
Reoperation rate	3 trials (n=625)	0.63 [0.32, 1.26]	1.31	0.19	2.43	0.30	18
Work status	3 trials (n=588)	1.18 [0.74, 1.86]	0.70	0.48	3.49	0.18	43
Surgery again	3 trials (n=514)	2.92 [1.34, 6.37]	2.70	0.007	5.54	0.06	64
Complications	4 trials (n=692)	0.50 [0.29, 0.84]	2.63	0.008	2.62	0.45	0

ODI = Oswestry disability index; VAS = Visual analog score; MD = Mean difference; OR = Odds ratios; CI = Confidence interval; NA = Not applicable.

ODI (Oswestry Disability Index) and VAS pain scores: The ODI low back pain disability questionnaire is a validated method of assessing a patient's level of pain and functional disability, and the VAS pain scores is used to assess the intensity and duration of back and leg pain. At 2 years, all patients showed significant improvement in ODI and VAS pain scores compared with baseline regardless of treatment. There were five trials reporting the two continuous outcomes (mean ± SD), and they were all included in the meta-analysis.[6,9,13,15,16] Five trials enrolled 1007 patients, with 667 patients being assigned into TDR group and the other 340 ones into fusion group. As for ODI and VAS pain scores, the test for heterogeneity revealed that there was no significant heterogeneity across the trials ($p=0.81$, $I^2=0\%$; $p=0.97$, $I^2=0\%$, respectively), and the fixed model was performed. Overall, TDR-treated patients showed a significant decrease in ODI scores ($p=0.001$) and VAS pain scores ($p=0.01$) compared to fusion-treated patients (Table IIIA, Figure 2).

Patient satisfaction or VAS patient satisfaction: The patient satisfaction questionnaire is a question on patient satisfaction with their treatment and a global outcome score of pain, and the VAS patient satisfaction is used to assess patients' satisfaction level with treatment by placing a mark on a printed 100-mm scale, with a higher score representing a better satisfaction. There were two trials reporting the dichotomous outcomes of patient satisfaction (OR),[6,15] and another two trials reporting the continuous outcomes of VAS patient satisfaction (mean ± SD).[13,16] The test for heterogeneity did not detect significant heterogeneity across the trials reporting patient satisfaction ($p=0.12$, $I^2=58\%$) and trials reporting VAS patient satisfaction ($p=0.92$, $I^2=0\%$). Using a fixed effects model, pooled results revealed that at 24 months TDR-treated patients had a significantly higher patient satisfaction ($p=0.002$) and VAS patient satisfaction ($p=0.002$) when compared with fusion-treated patients (Table IIIA, Figure 2).

Narcotic use: Data required for this meta-analysis was available from three trials.[6,13,16] Regardless of treatment, the percentage of patients reported use of narcotics to control pain was significantly decreased at 2 years compared to before surgery. Totally, the percentage of patients using narcotics in the TDR group was 56.3% (287/510) and fusion group was 69.5% (162/233). The test for heterogeneity revealed that there was a significant heterogeneity across the trials ($p=0.07$, $I^2=61\%$), and thus a random effects model was performed. Pooled results showed that TDR-treated patients had a significant lower percentage of using narcotics ($p=0.03$) compared to fusion-treated patients (Table IIIA, Figure 2).

Overall success rate: To be considered as an overall success, patients have to achieve all of the following: a 25% improvement in ODI score at 24 months compared with the pre-operative score, no device failure, no major complications, and no neurological deterioration compared to pre-operative status. The overall success rate is defined as the percentage of individual patients

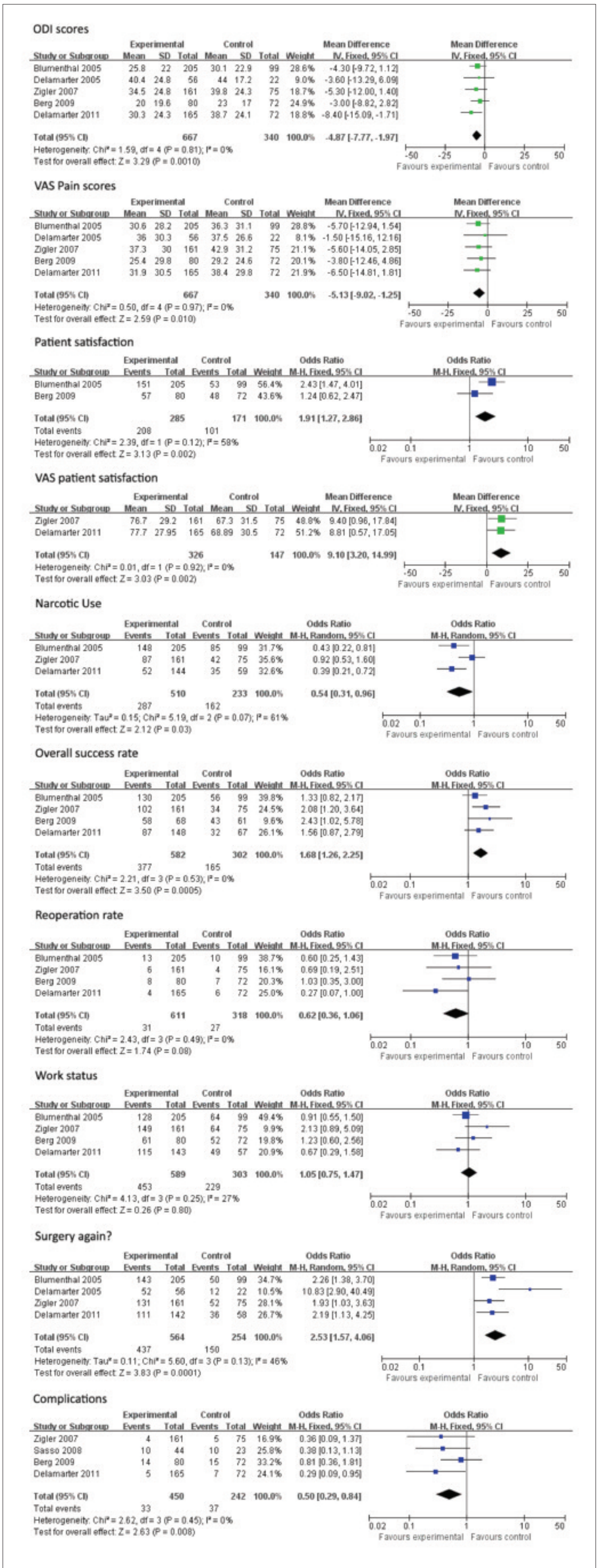

Figure 2: Pooled results of Total Disc Replacement (TDR) versus lumbar fusion at 2 years follow-up.

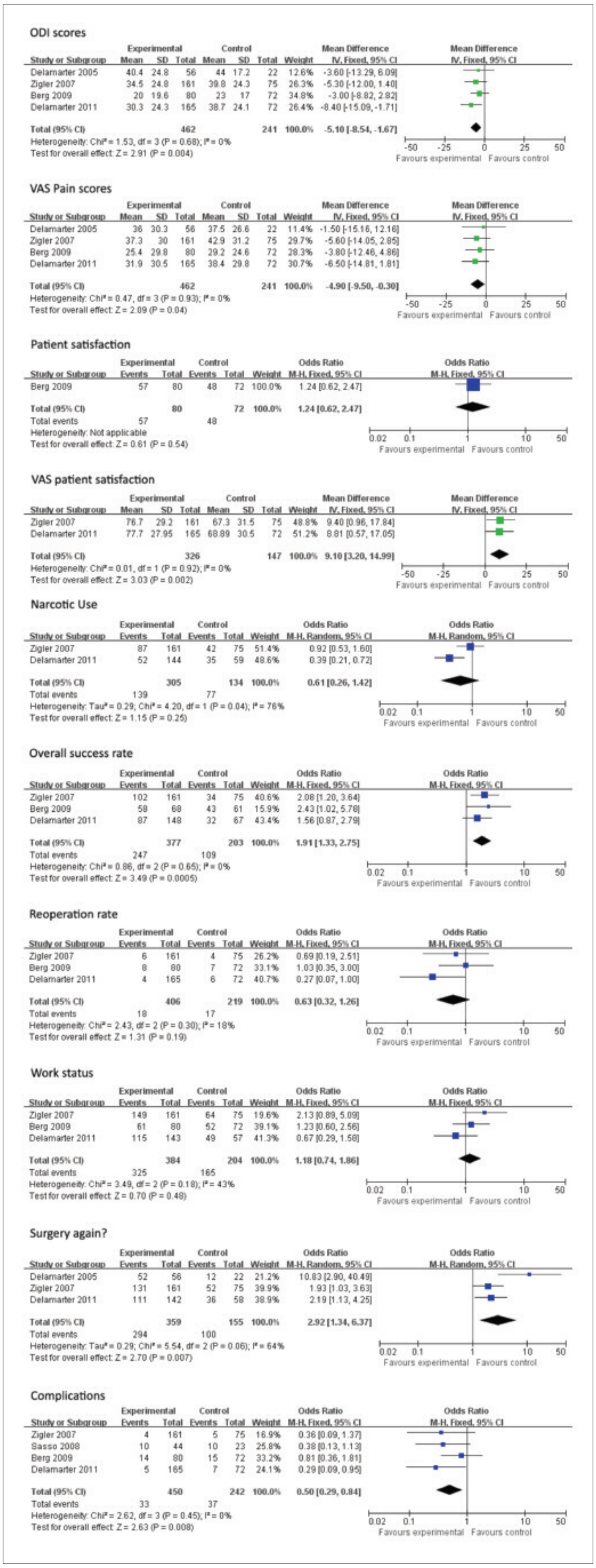

Figure 3: Pooled results of TDR versus lumbar fusion after excluding the study with stand-alone cage interbody fusion.

achieving success in all four-component criteria and is reported in four trials.[6,13,15,16] Totally, the overall success rate was 64.8% (377/582) in the TDR group and 54.6% (165/302) in the fusion group. The test for heterogeneity demonstrated that there was no significant heterogeneity across the four studies (p=0.53, I^2=0%). And pooled results in a fixed effects model showed that there was a significant increase in the overall success rate of TDR-treated patients compared to fusion-treated patients (p=0.005, Table IIIA, Figure 2).

Reoperation rate: Secondary surgical procedures, defined as any revision, removal, or reoperation of the implant or supplemental fixation, were recorded in four trials.[6,13,15,16] At 2 years, the rate of patients with secondary surgical procedures in the TDR group was 5.1% (31/611) and the fusion group was 8.5% (27/318). The test for heterogeneity demonstrated that there was no significant heterogeneity across the four studies (p=0.49, I^2=0%), and the fixed model was performed. Overall, pooled results showed that there was no significant difference regarding the rate of patients with secondary surgical procedures between the TDR group and fusion group (p=0.08). The details are shown in Table IIIA and Figure 2.

Work status: Work status refers to the percentage of patients partaking in the work both full and part-time and is investigated in four trials.[6,13,15,16] According to our meta-analysis, 453 and 229 patients [who received TDR (n=589) and fusion (n=303) respectively] were back at work (full or part-time) at 24 months. The test for heterogeneity demonstrated that there was no significant heterogeneity across the four studies (p=0.25, I^2=27%), and the fixed model was performed. Overall, pooled results revealed that TDR-treated patients did not have a significant higher percentage of employment (p=0.80) compared to fusion-treated patients. The details are shown in Table IIIA and Figure 2.

Surgery again? Four trials reported the responses of patients whether they would have the same surgical treatment again and were all included in the meta-analysis.[6,9,13,16] Overall, the percentage of patients responded “yes” at 2 years in the TDR group and fusion group was 77.5% (437 of 564) and 59.1% (150 of 254), respectively. The test for heterogeneity demonstrated that there was no significant heterogeneity across the four studies (p=0.13, I^2=46%), and the fixed model was performed. Pooled results showed that there was a significant higher rate of patients to chose the same surgical treatment again in the TDR group compared to fusion group (p=0.001) (Table IIIA, Figure 2).

Complications: The complications are the composite of major complications (major vessel injury, neurologic damage, nerve root injury, death and so on) and minor complications (clinically significant blood loss, retrograde ejaculation, infections, deep venous thrombosis, etc.). Overall, the complications were recorded in four trials,[13-16] and occurred in 33 of the 450 patients (7.3%) in the TDR group, as compared with 37 of the 242 patients (15.3%) in the fusion group. Pooled results in a fixed effects model suggested that the incidence of complications was significantly lower in the TDR group than in the fusion group (p=0.008, Table IIIA, Figure 2), and there was no heterogeneity among the studies (p=0.45, I^2=0%).

Radiographic outcomes: The radiographic outcomes that were assessed mainly included the Range of Motion (ROM) and the disc height. ROM at the 2-year postoperative follow-up was reported in three trials. In the trial by Blumenthal *et al.*,[6] ROM in the TDR group had an increase of 13.6% compared with pre-operative ROM, while mean ROM decreased as expected in the fusion group (averaged 1.1°). In the trial by Zigler *et al.*[13], 93.7% of TDR-treated patients had a normal functional range of ROM (averaged 7.7°). In the trial by Berg *et al.*,[15] ROM in the TDR group had increased compared with pre-operative values, which was mainly in the extension domain. Disk height at 24 months postoperatively was reported in two trials. In the trial by Blumenthal *et al.*,[6] TDR was significantly more effective than fusion for restoring the height of collapsed disc ($p < 0.05$). In the trial by Berg *et al.*,[15] disc height was still less than normal after fusion, while after TDR disc height was higher than normal (+2 SD). Moreover, there was a significant difference regarding postoperative disc heights between the two groups at 24 months ($p < 0.001$). These above results indicate that TDR not only results in increasing in the ROM but also restoring the disc height, which will reduce the biomechanical changes associated with the fusion.

DISCUSSION

Although many studies suggest that the effects of TDR for patients with symptomatic, single-level lumbar disc disease were equivalent or superior to those of lumbar fusion at 2 years of follow-up, relatively few reviews have comprehensively compared the long-term clinical outcomes of TDR with lumbar fusion for the treatment of patients with DDD in terms of meta-analysis.

A previous meta-analysis showed that compared to lumbar fusion, TDR results in a slightly better ODI and VAS pain scores and a significantly greater patient satisfaction as well as a significantly higher rate of patients would have the same surgical treatment again at the 2-year follow-up.[22] As for the complication, reoperation rate and patients' work status, no significant difference was detected between the two groups at 24 months. But when one study was excluded due to the fusion technique had a high influence on the overall results,[6] there was no significant difference at any of the above efficacy endpoints, which led to the conclusion that TDR does not show significant superiority for the treatment of lumbar DDD compared with fusion. However, due to the small number of eligible studies included in the meta-analysis especially when one of the RCTs was excluded, the validity of these results needs further confirmation and the conclusion of this review is not convincing.

Therefore, in this meta-analysis, we included more RCTs with upto 2 years follow-up in order to comprehensively compare the long-term clinical outcomes of TDR with lumbar fusion for the treatment of patients with DDD. Our results demonstrated that when compared to lumbar fusion, TDR yields better clinical outcomes regarding the ODI and VAS pain scores, patient satisfaction or VAS patient satisfaction, the rate of narcotic usage for pain, overall success rate, the rate of patients would have the same surgical treatment again and complications at the 2 years follow-up (Table IIIA, Figure 2). As for the reoperation rate and patients' work status, there is no significant difference between the two groups at 24 months (Table IIIA, Figure 2). The heterogeneity across these trials was slight, so most evidences from this study should be considered to be robust. However, significant heterogeneity was detected in the analysis of narcotic usage (p=0.07, I^2=61%), and the major contributor to the heterogeneity was the study by Zigler *et al*.[13] By removing this study, the heterogeneity was eliminated.

According to our meta-analysis, the pooled results of most efficacy endpoints are consistent with the previous meta-analysis except the complications. However, when the study by Blumenthal *et al*.[6] was excluded with the same reason in this meta-analysis, the pooled results showed that there was still significant difference regarding the ODI and VAS pain scores, VAS patient satisfaction, overall success rate, the rate of patients would have the same surgical treatment again and complications between the two groups at 24 months (Table IIIB, Figure 3). While for the outcomes of patient satisfaction, the rate of narcotic usage for pain, reoperation rate and patients' work status, the pooled results showed that there is no significant difference between the two groups at 24 months. The details are shown in Table IIIB and Figure 3. Overall, the above analysis may lead to the conclusion that TDR yields better long-term clinical outcomes for the treatment of lumbar DDD compared with fusion.

Nevertheless, this study still has several potential limitations. One potential limitation is that the types of the disc prosthesis and the control intervention of the included trials are not completely consistent, and this might cause a bias. A second potential limitation involves the fact that based on the generation of allocation sequence, allocation concealment, blinding, and the follow-up, most included trials were considered to be of low methodological quality due to lack of two or more unclear or inadequate quality components. A third confounder is that the small sample size of all included trails. A fourth potential limitation is that the number of eligible studies is still not enough, and the validity of the results needs more RCTs for further confirmation. In addition, even though the number of included trials for this meta-analysis is relatively small and a funnel plot for pooled estimates is not performed, there may be publication bias as well.

CONCLUSION

The present meta-analysis of RCTs reveals that in a long-term of follow-up (2 years) TDR shows a significant superiority for the treatment of lumbar DDD compared with fusion. However, due to the number of eligible studies in this meta-analysis which are still not enough, more high quality RCTs with a long-term follow-up (at least 2 years) are further needed to confirm the clinical benefits with the use of TDR in treatment of lumbar DDD.

REFERENCES

1. Hochschuler SH, Ohnmeiss DD, Guyer RD, Blumenthal SL. Artificial disc: preliminary results of a prospective study in the United States. *Eur Spine J* 2002; **11**:S106-10.
2. Frymoyer JW, editor. The adult spine: principles and practice. 2nd ed. Philadelphia: *Lippincott-Raven*; 1997.
3. Fritzell P, Hagg O, Wessberg P, Nordwall A. 2001 Volvo Award Winner in Clinical Studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. *Spine* 2001; **26**:2521-32; discussion 32-4.
4. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Hanscom B, Skinner JS, *et al.* Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. *JAMA* 2006; **296**:2441-50.
5. Gibson JN, Waddell G. Surgery for degenerative lumbar spondylosis: updated cochrane review. *Spine* 2005; **30**:2312-20.
6. Blumenthal S, McAfee PC, Guyer RD, Hochschuler SH, Geisler FH, Holt RT, *et al.* A prospective, randomized, multicenter food and drug administration investigational device exemptions study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part I: evaluation of clinical outcomes. *Spine* 2005; **30**:1565-75; discussion E387-91.
7. Ghiselli G, Wang JC, Bhatia NN, Hsu WK, Dawson EG. Adjacent segment degeneration in the lumbar spine. *J Bone Joint Surg Am* 2004; **86**-A:1497-503.
8. Akamaru T, Kawahara N, Tim Yoon S, Minamide A, Su Kim K, Tomita K, *et al.* Adjacent segment motion after a simulated lumbar fusion in different sagittal alignments: a biomechanical analysis. *Spine* 2003; **28**:1560-6.
9. Delamarter RB, Bae HW, Pradhan BB. Clinical results of ProDisc-II lumbar total disc replacement: report from the United States clinical trial. *Orthop Clin North Am* 2005; **36**: 301-13.

10. Bono CM, Garfin SR. History and evolution of disc replacement. *Spine J* 2004; **4**:145S-50S.

11. Le Huec J, Basso Y, Mathews H, Mehbod A, Aunoble S, Friesem T, *et al.* The effect of single-level, total disc arthroplasty on sagittal balance parameters: a prospective study. *Eur Spine J* 2005; **14**:480-6.

12. Dmitriev AE, Gill NW, Kuklo TR, Rosner MK. Effect of multilevel lumbar disc arthroplasty on the operative and adjacent-level kinematics and intradiscal pressures: an *in vitro* human cadaveric assessment. *Spine J* 2008; **8**:918-25.

13. Zigler J, Delamarter R, Spivak JM, Linovitz RJ, Danielson GO, Haider TT, *et al.* Results of the prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement versus circumferential fusion for the treatment of 1-level degenerative disc disease. *Spine* 2007; **32**:1155-62; discussion 63.

14. Sasso RC, Foulk DM, Hahn M. Prospective, randomized trial of metal-on-metal artificial lumbar disc replacement: initial results for treatment of discogenic pain. *Spine* 2008; **33**:123-31.

15. Berg S, Tullberg T, Branth B, Olerud C, Tropp H. Total disc replacement compared to lumbar fusion: a randomised controlled trial with 2-year follow-up. *Eur Spine J* 2009; **18**: 1512-9.

16. Delamarter R, Zigler JE, Balderston RA, Cammisa FP, Goldstein JA, Spivak JM. Prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement compared with circumferential arthrodesis for the treatment of two-level lumbar degenerative disc disease: results at twenty-four months. *J Bone Joint Surg Am* 2011; **93**:705-15.

17. Shea B, Dubé C, Moher D. Assessing the quality of reports of systematic reviews: the Quorum statement compared to other tools. In: Egger M, Smith GD, Altman DG, editors. Systematic reviews in healthcare: meta-analysis in context. London: *BMJ Publishing Group*; 2006.p. 122-39.

18. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; **273**:408-12.

19. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, *et al.* Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 1998; **352**:609-13.

20. Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med* 2001; **135**:982-9.

21. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**:629-34.

22. Yajun W, Yue Z, Xiuxin H, Cui C. A meta-analysis of artificial total disc replacement versus fusion for lumbar degenerative disc disease. *Eur Spine J* 2010; **19**:1250-61.