



Original Article

Corresponding Author

Jun Ho Lee

<https://orcid.org/0000-0001-6246-8053>

Department of Neurosurgery, Kyung Hee University Medical Center, 23 Kyungheedaero, Dongdaemun-gu, Seoul 02447, Korea
Email: moo9924@khu.ac.kr

Received: May 3, 2023

Revised: July 12, 2023

Accepted: July 15, 2023

Clinical Effectiveness of Artificial Disc Replacement in Comparison With Anterior Cervical Discectomy and Fusion in the Patients With Cervical Myelopathy: Systematic Review and Meta-analysis

Jung Hwan Lee¹, Youn Joo Lee², Min Cheol Chang³, Jun Ho Lee⁴

¹Department of Rehabilitation Medicine, Namdarun Rehabilitation Clinic, Yongin, Korea

²Chadwick International School, Incheon, Korea

³Department of Physical Medicine and Rehabilitation, College of Medicine, Yeungnam University, Daegu, Korea

⁴Department of Neurosurgery, Kyung Hee University Medical Center, Seoul, Korea

Objective: Cervical myelopathy (CM) describes the compressive cervical spinal cord state, often accompanied by serious clinical condition, by herniated disc or hypertrophied spurs or ligament. Anterior cervical discectomy and fusion (ACDF) has been frequently employed as conventional surgical solution for this CM despite its inherent biomechanical handicap. Alternatively, an artificial disc replacement (ADR) preserves cervical motion while still decompressing the spinal canal and neural foramen. This analysis elaborated to clarify the potential benefits of ADR application to CM over ACDF from the conglomerated results of the past references.

Methods: A literature search was performed using MEDLINE, Embase, Cochrane review, and KMBase databases from the studies published until March 2023. Six studies (3 randomized controlled study [RCTs] and 3 non-RCTs) were included in a qualitative and quantitative synthesis. Data were extracted and analyzed using a random effects model to obtain effect size and its statistical significance. Quality assessment and evidence level were established in accordance with the GRADE (Grading of Recommendations Assessment, Development and Evaluation) methodology.

Results: Among 6 studies, 2 studies showed that ADR group achieved significantly better clinical improvement than the ACDF group, while the rest 4 studies revealed no significant difference. A meta-analysis showed better clinical outcomes with or without statistical significance. The level of evidence was low because of inconsistency and imprecision.

Conclusion: ADR was superior or at least, not inferior to ACDF in terms of functional recovery. However, its application to the CM patients is merely empowered with weak strength due to low level of evidence.

Keywords: Cervical myelopathy, Anterior cervical discectomy and fusion, Artificial disc replacement, Clinical outcome, Systematic review, Meta-analysis



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2023 by the Korean Spinal Neurosurgery Society

INTRODUCTION

Cervical myelopathy (CM) is a condition where the cervical

spinal cord is compressed either ventrally by herniated disc or dorsally by hypertrophied bony structures or ligament, frequently yielding the serious clinical manifestations ranging

from neck and/or arm pain, functional deficits limited to upper limb involvement even to lower limb functional impairment including gait disturbance.¹⁻³ Conservative managements are frequently fraught with failure during its treatment, subsequently requesting an early, surgically decompressive management when the patients manifest excessive intractable pain or progressive neurological deficits.

Anterior cervical discectomy and fusion (ACDF) has been playing its role as an effective solution for CM.^{4,5} It might provide the sufficient decompression as well as the solid stabilization to the affected cervical segment. However, few concerns that relate to its inherent potency such as fusing the mobile spinal segment as well as incurring subsequent adjacent segment degeneration (ASD)^{6,7} have brought up the artificial disc replacement (ADR) as an alternative. The ADR might offer both the mechanical advantage of segmental motion preservation with consequent stress reduction at the adjacent levels while fulfilling the sufficient cervical cord decompression as during the ACDF.⁸⁻¹⁰ But there have been some concerns whether this mobile ADR can yield compatible clinical outcome and eventually be a substitute to ACDF in the treatment of CM. Past references have noticed the existence of disparities in the application of ADR versus ACDF in terms of the socioeconomic aspect, that the subjects with greater median income or the possession of the private insurance preferred ADR.¹¹ They also manifested the elaborations to minimize its application during the surgical treatment of multilevel CM by overlapping with laminoplasty.¹² Taking into account that the majority of the CM subjects are tend to be senile, deformed populations with already progressed facet degeneration as well as decreased cervical lordosis from even spontaneous fusion, ACDF still comprises as a general surgical solution for CM over ADR.¹³

There have been a few clinical studies that compare clinical results between ADR and ACDF in the CM treatment. Hereby, we conducted systemic review with meta-analysis by synthesizing published articles regarding this topic to clarify the clinical benefits of ADR over the ACDF and further to investigate its potential as a substitute to ACDF for the CM management.

MATERIALS AND METHODS

1. Study Selection Criteria

The authors have recruited articles described in Korean or English language that have primarily met the following criteria: patients aged from 18 to 70 years old, clinical manifestation of CM with no significant improvement or even aggravation after

conservative treatment, and the confirmative diagnosis of 1 or 2 levels of mechanical cervical cord compression by computed tomography (CT) or magnetic resonance imaging (MRI). Exclusion criteria were a previous history of cervical spinal surgery, severe osteoporosis, cervical kyphosis, foraminal stenosis, severe facet joint spondylosis, cervical radiculopathy, ossification of the posterior longitudinal ligament, inflammatory cervical cord diseases, tumor, or infectious disease. Among the studies fulfilling these criteria, those that have included the contents regarding the clinical outcome after the ADR or ACDF and have provided the comparative results between the 2 surgical methods were finally selected.

2. Database Search and Study Extraction

The MEDLINE (PubMed), Embase, Cochrane review, and KMBase databases were searched for articles published until March 2023. We established individual search terms in each database's search engine (Supplementary Material). The search was not restricted to randomized controlled study (RCT) and was extended to original articles, including non-RCT. The decision for an article selection was primarily based on the title and abstract review, followed by full-text screening. The study screening and data extraction were independently performed by the 2 reviewers, and any discrepancies were resolved by discussion between the 2 reviewers or with the entire research group. Flow chart demonstrating the process of study selection was illustrated in Fig. 1.

3. Data Collection

Reference data such as the diagnosis and number of subjects in each surgical method group (ADR and ACDF groups), clinical evaluation tools, follow-up period, and comparative results of the clinical outcomes were extracted from the selected articles. Dichotomous variables such as the number of patients with pain and functional scores or adverse events were extracted for the estimation of relative risk ratio. Continuous variables such as mean and standard deviation of clinical scores were extracted for the estimation of mean differences. If the standard deviations were not reported, they were calculated from confidence interval (CI), mean, and the number of patients.

4. Quality Assessment of Selected Studies, Establishment of Level of Evidence, and Strength of Recommendation

Quality assessment of each study and level of evidence was established in accordance with the Grading of Recommendations Assessment, Development and Evaluation (GRADE)

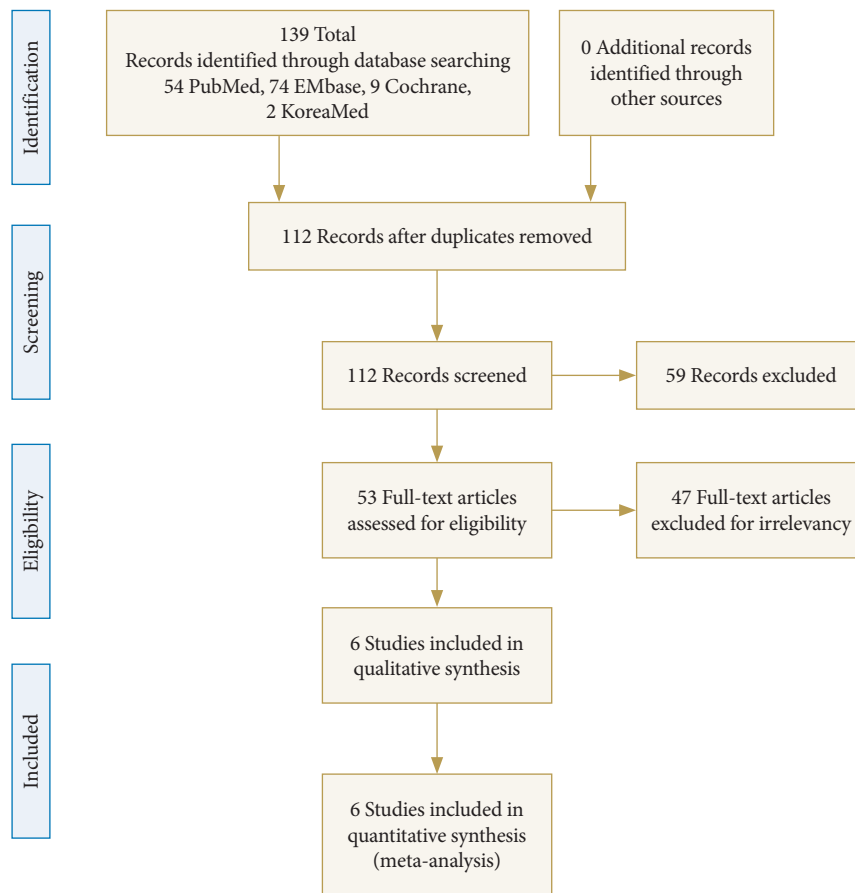


Fig. 1. Flow diagram of study selection.

methodology.^{14,15} The bias assessment for each RCT was conducted by method of risk of bias (ROB), which consisted of 7 domains: random sequence generation, allocation sequence concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective reporting, and other biases.¹⁶ The bias for each non-RCT was assessed with Risk of Bias Assessment tool for Nonrandomized Study (RoBANS); domains were selection of participants, confounding variables, measurement of intervention (exposure), blinding of outcome assessment, incomplete outcome data, and selective reporting.¹⁷ All the domains were evaluated as “low risk,” “high risk,” or “unclear.” These evaluations were performed by 2 independent reviewers and disagreements were resolved by discussion between 2 reviewers or with the entire research group.

Based on the comprehensive evaluation of inconsistency, indirectness, imprecision and publication bias in addition to ROB in all studies, the evidence level was determined as high, moderate, low, or very low grade. Besides, the strength of recommendation was determined as strong or weak by comprehensively

assessing not only evidence level, but also other factors such as benefits, risks, burdens, and possibly cost.¹⁸ The level of evidence and strength of recommendation were determined by discussion involving the entire research group.

5. Meta-analysis

Review Manager software (RevMan ver. 5.4; Cochrane Collaboration, Oxford, UK) was used for meta-analysis to compare clinical outcome between the ADR and the ACDF group. Tests of heterogeneity were performed using I^2 statistics. The parameter with I^2 values of $p < 0.05$, which was considered to have significantly high degree of heterogeneity, was additionally validated by subgroup analysis or sensitivity analysis. A random effects model was applied to obtain effect size and its statistical significance because it was assumed that the subjects and methods of the included studies performed by independent researchers could not be entirely equivalent and, therefore, could not have a common effect size. A probability of $p < 0.05$ was considered statistically significant. The results were expressed as mean difference and 95% CI for continuous outcome data

and in the form of relative risk ratio and 95% CI for dichotomous outcome data.

RESULTS

1. Search Results

Our database search has initially recruited 139 articles. After the exclusion of the 27 duplicates, 112 potentially eligible articles have remained. After the title and abstract screening, 59 articles were excluded due to the lack of the inclusion criteria fulfillment. Thus, the remaining 53 articles were retrieved for full-text analysis, of which 47 were subsequently excluded because of the irrelevance to the scheme of this analysis.

Ultimately, 3 RCTs and 3 non-RCTs were included in this study (Fig. 1).¹⁹⁻²⁴ The pain intensity was measured in the selected studies using Numerical Rating Scale or visual analogue scale. The functional measurement tool used in the selected studies was the Neck Disability Index (NDI), Japanese Orthopaedic Association score (JOA), and 36-item Short Form health survey (SF-36). Odom criteria were used to evaluate the degree of patients' satisfaction for treatment, which was divided into 4 grades including excellent, good, fair and poor grades. Excellent or good grade was regarded as satisfactory response. The follow-up period was variable across the studies ranging from 3 weeks to 7 years.

2. Quality Assessment

The ROB of all selected studies was illustrated in Fig. 2 (A: RCT, B: non-RCT). Three RCTs were assessed as unclear risk in random sequence generation domain because they did not describe the sequence generation process.^{19,22,24} The most frequently biased domain was blinding of outcome assessment, in which all 3 RCTs were rated as high risk or unclear because one RCT assessed the patients by nonblind staff,¹⁹ and the other 2 studies did not clarify whether the clinical evaluation was conducted by assessor who was blind or not involved in the process of patients selection and treatment.^{22,24} Sixteen domains among 21 (76.2%) were rated as low risk, thus, the overall ROB was considered low (Fig. 2A). Of 3 non-RCTs, one study was rated as high risk in selection of participants domain because they compared the group consisting of the patients recruited during different period.²¹ All 3 non-RCTs did not reveal whether clinical outcome was evaluated by the staff blind to treatment and thus were rated as unclear risk.^{20,21,23} Of 18 domains across all studies, 13 domains (72.2%) were determined as low risk; thus, the overall ROB was considered low (Fig. 2B). A discrepancy between reviewers was found in 6 of total 39 domains (15.4%) at first. After the discussion, all the discrepancies were resolved.

3. Clinical Outcome Analysis

All included studies have disclosed significant improvement of clinical results after 2 types of surgery. Among the 3 RCTs ul-

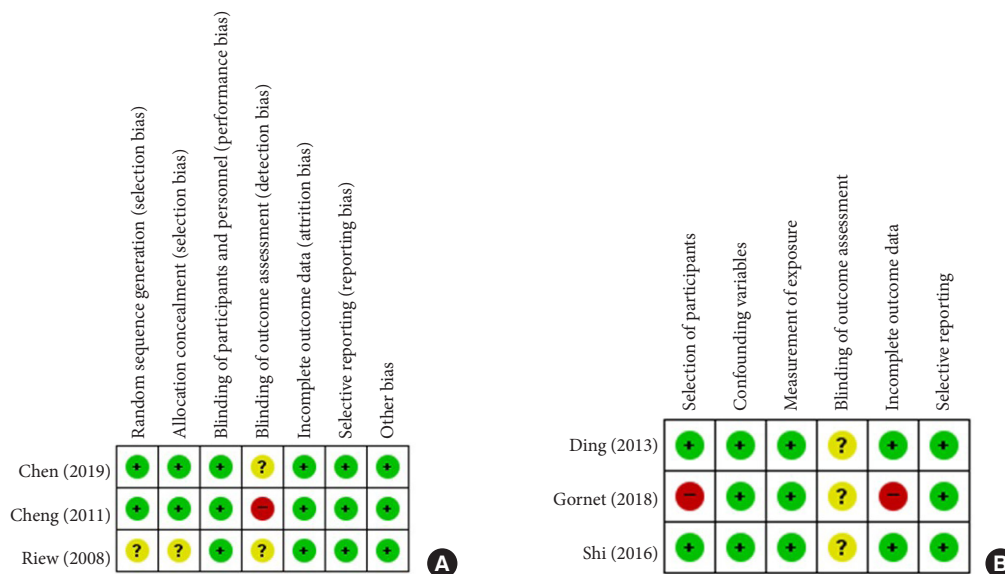


Fig. 2. Quality assessment for extracted studies. (A) Risk of bias for randomized controlled study. (B) Risk of Bias Assessment tool for Nonrandomized Study for nonrandomized study. Green color, low risk of bias; red color, high risk; white color, unclear risk of bias.

timately selected, 1 study showed that the ADR group achieved significant better clinical improvement than the ACDF group,¹⁹ while the other 2 studies revealed no significant difference between the 2 groups.^{22,24} Of 3 non-RCTs, 1 study found better clinical results in the ADR group than the ACDF group,²⁰ whereas the other 2 studies found no significant different clinical outcomes between the 2 groups.^{21,23} Comprehensively, ADR was superior or at least noninferior surgical method to obtain favorable clinical outcomes in the patients with CM (Table 1).

4. Meta-analysis

Meta-analysis was mainly performed in terms of surgical time (minutes), blood loss (mL), number of reoperations, range of motion (ROM) of surgical level, functional scores, and patients' satisfaction score, sufficiently provided for analysis across the studies. The study by Riew et al.²² divided the ADR group into Prestige ST and Bryan group according to prosthesis type and compared them with the 2 ACDF group respectively. Thus, 2 comparisons were extracted and analyzed respectively in this meta-analysis. Ultimately 6 studies and 7 comparisons were included in meta-analysis. Because the data regarding pain score

was not provided sufficiently for a proper meta-analysis performance, this analysis was performed mainly as to the functional improvement after surgery. Since the preoperative baseline NDIs and JOAs were not consistent across the selected studies, the changes of mean and standard deviation between baseline and follow-up period was obtained and analyzed.

5. Surgical Time

Seven comparisons from 6 studies provided continuous data of surgical time for the analysis of effect size by the mean difference.¹⁹⁻²⁴ The overall mean difference was estimated as 13.74 (95% CI, 0.71–26.76), which meant that ACDF required longer surgical time with statistical significance ($p=0.04$). A high degree of heterogeneity was revealed ($I^2=93%$) ($p<0.01$) (Fig. 3A). Sensitivity test revealed no specific study to significantly contribute to overall heterogeneity.

6. Blood Loss

Seven comparisons from 6 studies provided continuous data of blood loss for the analysis of effect size by the mean difference.¹⁹⁻²⁴ The overall mean difference was estimated as -18.44

Table 1. Evidence table

Study	Design	Intervention	Evaluation	Results
Riew et al. ²² (2008)	RCT	N = 106 ADR (N = 59, Prestige ST, N = 47, Bryan) N = 93 ACDF	NDI, SF-36, neck and arm pain score at 6 weeks, 3, 6, 12, and 24 months	Both groups had significant improvement following surgery, which was not significantly different between the groups.
Cheng et al. ¹⁹ (2011)	RCT	N = 41 ADR (Bryan) N = 42 ACDF	Modified Odom's criteria, JOA scale, SF-36, and NDI at 3 years	The ADR group scored significantly better in JOA scale, SF-36, and NDI than the ACDF.
Chen et al. ²⁴ (2019)	RCT	N = 30 ADR (Bryan) N = 30 ACDF	NDI, VAS, JOA scale, and Odom's criteria at 2 weeks, 3 months, 1 year, and 3 years	There were no significant differences in VAS, JOA, NDI, and satisfactory rates of Odom criteria between the 2 groups before or after the operation.
Ding et al. ²⁰ (2013)	Retrospective comparative study	N = 37 ADR (N = 25, Prestige ST, N = 12, Bryan) N = 39 ACDF	SF-36 for PCS and MCS, JOA scale, NDI, Nurick grade at 1 week, 3 months, 6 months, 12 months, and 24 months	The ADR group achieved significant improvement in NDI, SF-36 for PCS and MCS than the ACDF group. Both group showed the significant improvement in JOA scale and Nurick grade, which was no significant difference between both groups.
Gornet et al. ²¹ (2018)	Retrospective comparative study	N = 59 ADR (Prestige ST) N = 51 ACDF	NDI, neck & arm pain, SF-36, neurological status, at 2 and 7 years	No significant differences in clinical outcomes were found between the 2 groups.
Shi et al. ²³ (2016)	Prospective cohort study	N = 60 ADR (Discover) N = 68 ACDF	JOA scale and NDI until 24 months	Both treatments significantly improved all clinical parameters without statistically relevant differences between both groups.

ADR, artificial disc replacement; ACDF, anterior cervical discectomy and fusion; NDI, Neck Disability Index; SF-36, 36-item Short Form health survey; JOA, Japanese Orthopaedic Association; VAS, visual analogue scale; PCS, physical composite score; MCS, mental composite score.

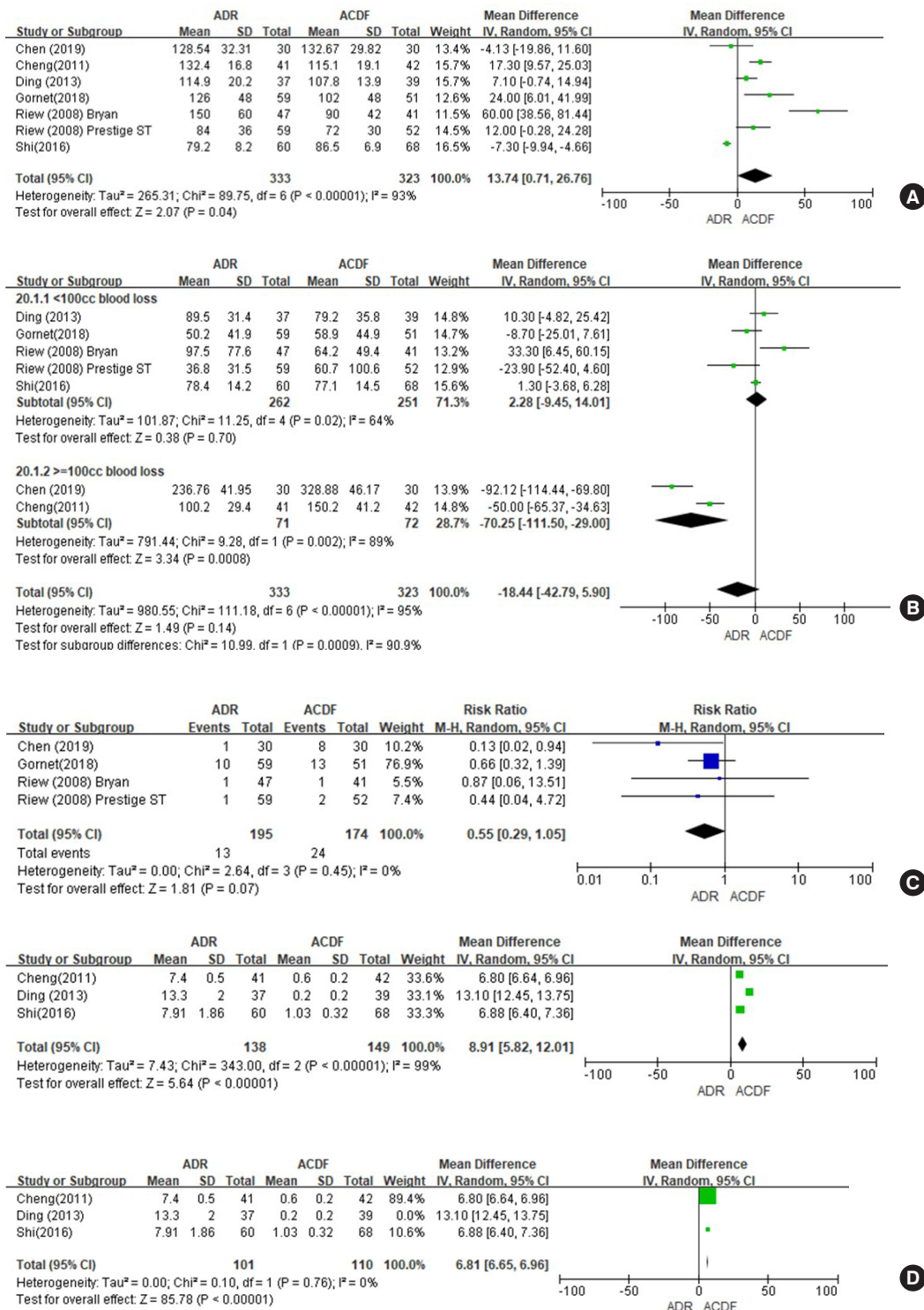


Fig. 3. Forest plot of comparison. (A) Mean difference of surgical time. (B) Mean difference of blood loss (upper: subgroup analysis [< 100 mL vs. ≥ 100 -mL blood loss], lower: total analysis). (C) Relative risk of reoperation. (D) Mean difference of range of motion (ROM) at surgical level (upper: total analysis, lower: sensitivity analysis after exclusion of the study of Ding et al.). Bryan, artificial disc replacement with Bryan prosthesis; Prestige ST, artificial disc replacement with Prestige ST; ADR, artificial disc replacement; ACDF, anterior cervical discectomy and fusion; SD, standard deviation; CI, confidence interval; df, degrees of freedom.

(95% CI, -42.79 to 5.9), meaning blood loss was smaller in ADR than ACDF, but without statistical significance ($p=0.14$). A high degree of heterogeneity was revealed ($I^2=95\%$) ($p<0.01$) (Fig. 3B).

Subgroup analysis was conducted after division of the studies into 2 subgroups depending on whether blood loss was 100 mL or more and less than 100 mL. In the subgroup of 100 mL or more of blood loss,^{19,24} the mean difference was estimated as -70.25 (95% CI, -111.5 to -29.0), meaning blood loss was significantly smaller in ADR than ACDF ($p<0.01$). But high degree of heterogeneity was also found, although its degree was somewhat decreased by subgroup analysis ($I^2=89\%$) ($p<0.01$). In the subgroup of less than 100 mL of blood loss,^{20,23} the mean difference was estimated as 2.28 (95% CI, -9.45 to 14.01), meaning blood loss was smaller in ACDF than ADR without statistical significance ($p=0.70$). But high degree of heterogeneity was also observed despite its decrease after subgroup analysis ($I^2=64\%$) ($p=0.02$) (Fig. 3B).

7. Reoperation

Four comparisons from 3 studies provided the dichotomous data for measurement of effect size by relative risk ratio about reoperation rate.^{21,22} This suggested that a smaller number of reoperations were observed in ADR than ACDF with estimated risk ratio of 0.55 (95% CI, 0.29–1.05) without statistical significance ($p=0.07$). No heterogeneity was found ($I^2=0\%$) ($p=0.45$) (Fig. 3C).

8. ROM at Surgical Level

Three studies were available in the analysis of effect size by the mean difference for ROM at surgical level.^{19,20,23} The estimated overall mean difference was calculated as 8.91 (95% CI, 5.82–12.01), which favored ADR with statistical significance ($p<0.01$). The degree of heterogeneity was significant ($I^2=99\%$) ($p<0.01$).

Sensitivity test without Ding et al.'s study revealed the low degree of heterogeneity ($I^2=0\%$) ($p=0.76$), which indicated that Ding et al.'s study significantly contributed to heterogeneity of ROM at surgical level. After removal of this study, the estimated overall mean difference was calculated as 6.81 (95% CI, 6.65–6.96), which revealed also favor of ADR with statistical significance ($p<0.01$) (Fig. 3D).

9. NDI at 3 Months

Five comparisons from 4 studies presented continuous data of NDI improvement at 3 months and were available in the analysis

of effect size by the mean difference.^{20,22,24} The overall mean difference was estimated as -3.52 (95% CI, -6.80 to -0.23), favoring ADR, and this was the degree of statistical significance ($p=0.04$). A high degree of heterogeneity was revealed ($I^2=82\%$) ($p<0.01$) (Fig. 4A). Sensitivity test revealed no specific study to significantly contribute to overall heterogeneity.

10. NDI at 6 Months

Three comparisons from 2 studies were available in the analysis of effect size by the mean difference for NDI improvement at 6 months.^{20,22} The estimated overall mean difference was calculated as -6.36 (95% CI, -12.26 to 0.46), which favored ADR with statistical significance ($p=0.03$). The degree of heterogeneity was not significant ($I^2=61\%$) ($p=0.08$) (Fig. 4B).

11. NDI at 12 Months

Four comparisons from 3 studies were available in the measurement of effect size by the mean difference of successful NDI improvement at 12 months.^{20,22,24} The estimated overall mean difference was calculated as -5.27 (95% CI, -9.79 to 0.75), which meant significant superiority of ADR over ACDF ($p=0.02$). The degree of heterogeneity was significantly high ($I^2=74\%$) ($p<0.01$).

Sensitivity test excluding Chen et al.'s study revealed the low degree of heterogeneity ($I^2=37\%$) ($p=0.20$), which indicated that this study significantly contributed to heterogeneity of NDI at 12 months. After removal of this study, the estimated overall mean difference was calculated as -7.18 (95% CI, -11.72 to -2.64), which empowered the significant superiority of ADR over the ACDF ($p<0.01$) (Fig. 4C).

12. NDI at 24 Months

Five comparative data from 4 studies provided the value of -5.26 (95% CI, -10.56 to 0.03), the effect size measured by the mean difference, which showed trends toward ADR without statistical significance ($p=0.05$).^{20,23} Significant heterogeneity was found ($I^2=89\%$) ($p<0.01$).

Sensitivity test excluding Shi et al.'s study revealed the low degree of heterogeneity ($I^2=43\%$) ($p=0.15$), which indicated that this study significantly contributed to heterogeneity of NDI at 24 months. The analysis after excluding this study obtained the statistical significance with the estimated overall mean difference of -7.09 (95% CI, -11.11 to -3.08), which favor ADR ($p<0.01$) (Fig. 4D).

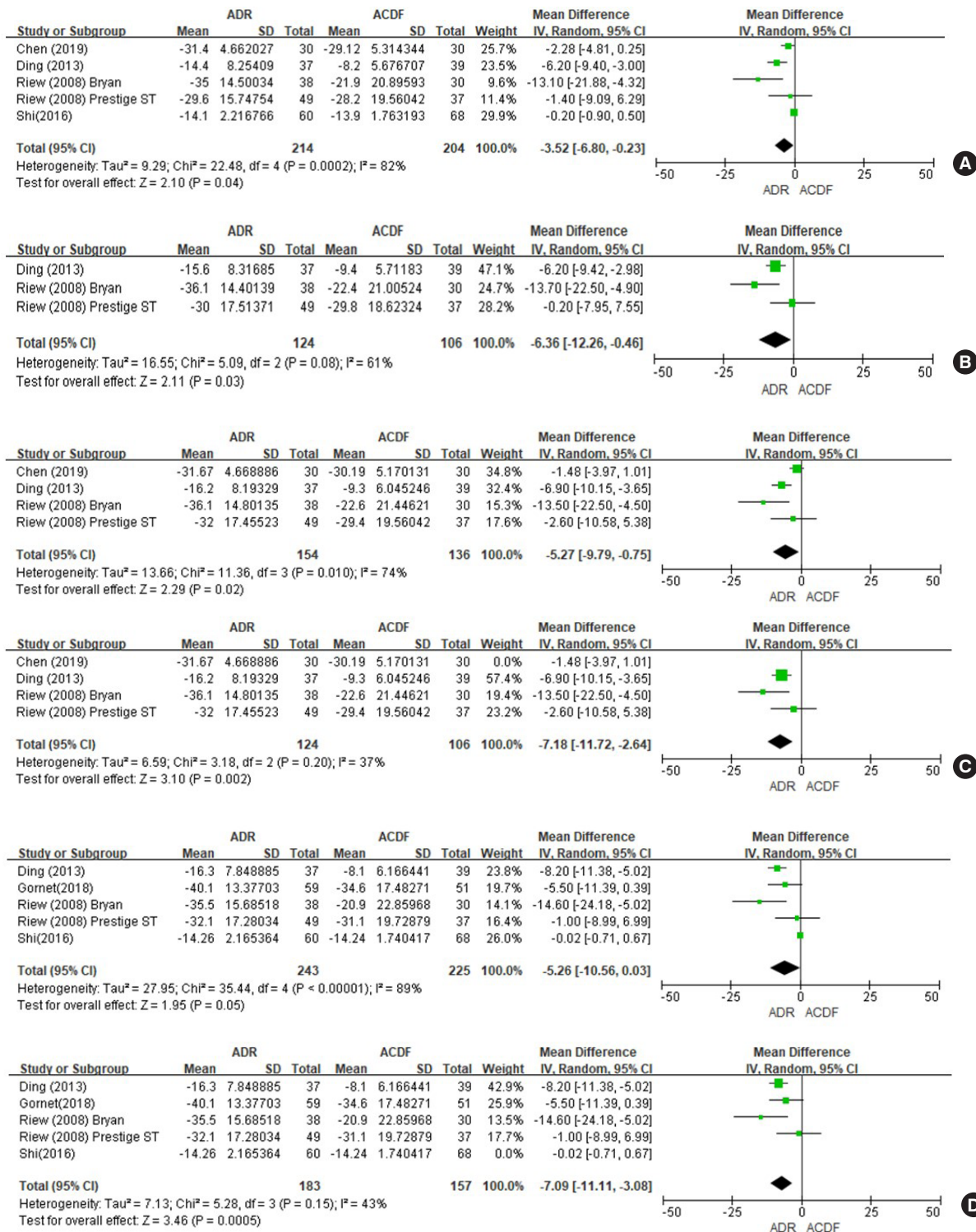


Fig. 4. Forest plot of comparison. (A) Mean difference of Neck Disability Score (NDI) change between the baseline and 3 months. (B) Mean difference of NDI change between the baseline and 6 months. (C) Mean difference of NDI change between the baseline and 12 months (upper: total analysis, lower: sensitivity analysis after exclusion of the study of Chen et al.). (D) Mean difference of NDI change between the baseline and 24 months (upper: total analysis, lower: sensitivity analysis after exclusion of the study of Shi et al.). Bryan, artificial disc replacement with Bryan prosthesis; Prestige ST, artificial disc replacement with Prestige ST; ADR, artificial disc replacement; ACDF, anterior cervical discectomy and fusion; SD, standard deviation; CI, confidence interval; df, degrees of freedom.

13. JOA at 3 Months

Two studies presented the continuous data for measurement of effect size by mean difference about JOA improvement at 3 months.^{23,24} The data showed no significance between ADR and ACDF with an estimated mean difference of 0.15 (95% CI, -0.27 to 0.57), despite slight supportiveness for ADR ($p = 0.48$). No heterogeneity was observed ($I^2 = 0\%$) ($p = 0.63$) (Fig. 5A).

14. JOA at 12 Months

Two studies were available in the measurement of effect size by mean difference of JOA improvement at 12 months.^{20,24} This analysis showed slightly favorable trends toward ADR with an estimated mean difference of 0.34 (95% CI, -0.78 to 1.46), but without statistical significance ($p = 0.55$). No significant heterogeneity was observed ($I^2 = 67\%$) ($p = 0.08$) (Fig. 5B).

15. JOA at 24 Months

Two studies provided the value of 0.23 (95% CI, -0.18 to 0.64), the effect size measured by the mean difference, which slightly favored ADR without statistical significance ($p = 0.28$).^{20,23} No significant heterogeneity was found ($I^2 = 0\%$) ($p = 0.67$) (Fig. 5C).

16. Patients' Satisfaction at 3 Years

Two studies provided the dichotomous data for measurement of effect size by relative risk ratio about JOA improvement at 3 years, which slightly favored ACDF with estimated risk ratio of 1.06 (95% CI, 0.91–1.23) without statistical significance ($p = 0.47$).^{19,24} No significant heterogeneity was found ($I^2 = 66\%$) ($p = 0.09$) (Fig. 5D).

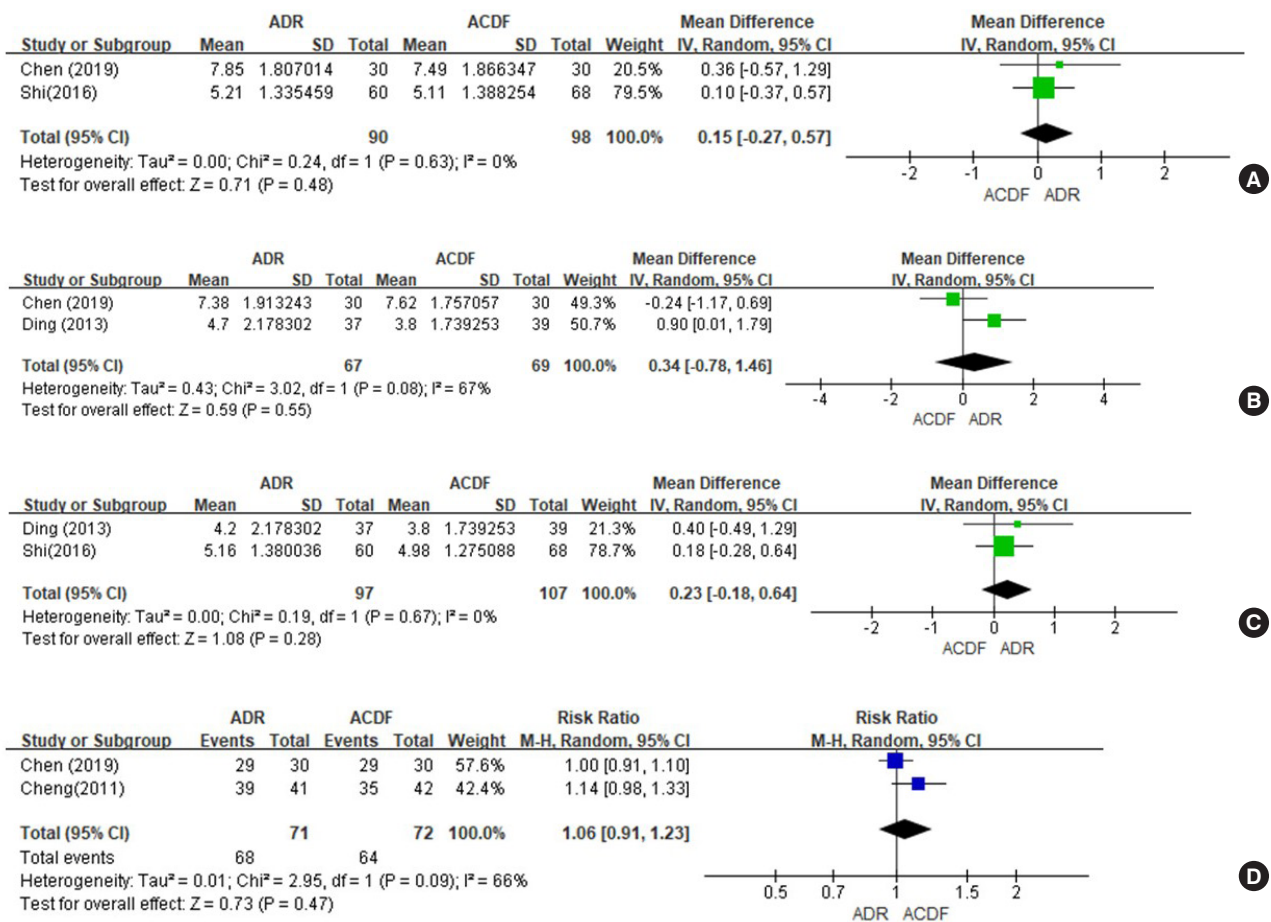


Fig. 5. Forest plot of comparison. (A) Mean difference of Japanese Orthopedic Association scale (JOA) change between the baseline and 3 months. (B) Mean difference of JOA change between the baseline and 12 months. (C) Mean difference of JOA change between the baseline and 24 months. (D) Relative risk ratio of patients' satisfaction at 3 years. ADR, artificial disc replacement; ACDF, anterior cervical discectomy and fusion; SD, standard deviation; CI, confidence interval; df, degrees of freedom.

17. Level of Evidence and Strength of Recommendation

The ROB was considered to be low (not-problematic or not serious) as previously mentioned. Directness was not considered as problematic because all included studies have directly compared ADR with ACDF. Publication bias was not assessed because fewer than 10 studies were included during each meta-analysis. The consistency was validated to be problematic due to clinical heterogeneity across the selected studies and statistical heterogeneity revealed by I^2 of some categories of meta-analysis. The degree of precision was also regarded as problematic or serious due to the small number of patients in the selected studies. Ultimately, evidence level determined by GRADE system is evaluated to be low.

Overall, synthetic results of selected studies and meta-analysis showed superior or at least, not inferior clinical outcomes of ADR over ACDF mainly in terms of the functional improvement during the treatment of CM. Although ADR might be costlier or demands more sophisticated surgical details, it might be superior over ACDF for the sake of cervical ROM preservation as well as ASD prevention in addition to the clinical benefits found in meta-analysis.

After all of these analyses and considerations, the authors have concluded that ADR could be recommended over ACDF for the surgical treatment of the patients with CM with weak strength.

DISCUSSION

CM is a form of myelopathy that involves compression of the spinal cord at the cervical level of the spinal column. The radiologically verified CM by MRI or CT scans often results in the clinical/symptomatic myelopathy of neurological deficit of spasticity (sustained muscle contractions), hyperreflexia, pathologic reflexes, loss of fine motor skills (digit/hand clumsiness), loss of balance, and/or subsequent gait disturbance. Usually, this condition prevails in the elderly population with the advanced cervical column degeneration after the repetitive wear-and-tear changes that inflicts over the cervical vertebrae as we age (cervical spondylotic myelopathy).

The benefits of ACDF during CM treatment would be its direct, cervical cord decompressing potential under less invasive operative field compared to the posterior approach, cervical alignment restorative capability, as well as stability provision. However, since the physiological curvature of the cervical spine mainly functions to reduce and buffer the external shock, the straightening or the reversal of the cervical lordosis combined

with immobilization of the single or a couple of the functional cervical segments after the ACDF procedures might sometimes propagate the tension of the paravertebral muscles and ligament complex, accelerates the degeneration of the adjacent cervical levels, and consequently incur pains in the neck muscles.²⁵

In this regard, the ADR has been accepted as an alternative to ACDF even for the CM populations recently since the same surgical approach and neural decompression is carried out as per ACDF but the motion segment is effectively preserved as well as minimize degenerative changes at adjacent levels as seen in ACDF. Although CM patients have been actually treated with ADR in previous studies, their results are usually mixed in with radiculopathy patients, making it difficult to assess their exclusive outcomes within a specific cohort.

This meta-analysis showed that ADR achieved compatible or better functional outcomes in terms of NDI and JOA in the treatment of CM from 3- to 24-month follow-up compared to ACDF. The NDI assesses the degree of interference of neck pain during daily activities. Therefore, NDI, although this might be the functional score, rather preferentially reflects the neck and upper limb pain severity. JOA evaluates the severity of functional aspect of upper and lower limb, reflecting neurological deficits resulted not only from peripheral nervous system but also from spinal cord or central nervous system. Consequently, the clinical benefits of ADR implementation would be supported in the pain relief, cervical motion maintenance as well as upper or lower limb functional recovery aspect.

A meta-analysis recently published also showed positivity of ADR compared with ACDF,² which compared clinical outcomes 2 surgical techniques in patients with CM, making it similar to the current study. But the current study has split the follow-up period from 3 to 24 months and has evaluated the clinical outcomes according to the differential time point, which might be the main differentiated aspect from the prior existing literature. The functional outcomes such as NDI and JOA between the 2 groups were compared after analyzing the improvement in terms of these functional scores rather than the final scores at the follow-up period. This helped to overcome the inconsistency of the preoperative baseline scores across selected studies and eventually further clarified the comparison. In addition, the current study has conducted the sensitivity test or subgroup analysis for the parameters with significantly high heterogeneity. If the outlier study was found, the authors have re-analyzed the parameter after the exclusion of that outlier, which might prompt the statistical robustness.

The few investigated parameters related to the operative pro-

cedure also added the superiority of ADR over the ACDF with or without statistical significance during this analysis. With these purported procedural simplicity as well as benefit such as shorter operative duration with less bleeding or reoperation rates, Hill et al.²⁶ has even recommended for a 2-level ADR in the outpatient setting for properly selected patients without increasing the complication or readmission rates.

The ADR group lost less blood during surgery than the ACDF group in this study, of which 4 studies indicated less blood loss of the ADR group in comparison with the ACDF group. Gao et al.²⁷ meta-analysis including the patients with radiculopathy, myelopathy, or disc herniation, which is the main difference from our study, showed the opposite results, explaining that increased blood loss was resulted from the keel cuts into cancellous bone required by arthroplasty technique. The main reason of opposite results observed in our study might be that ACDF required more extensive resection in association with large blood loss than ADR in case of myelopathy.

The longer surgical time of ADR than ACDF could be explained by that ADR required more skillful and exact placement of prosthesis to establish physiological axis of rotation, which was essential component of successful surgical results.^{28,29}

Reoperation rate revealed by 3 studies included in this analysis was about 6.67%. This was higher than the previously reported revision rate of 3.9% by another systemic review. This discrepancy might be incurred from lack of the clinical or radiological definition or indication on the prosthesis failure requesting the reoperation across the past references.³⁰

The main concern or controversy with ADR as the surgical choice for CM would be the feasible, repetitive microtrauma that might constantly inflicts on the spinal cord after the ROM maintenance at the ADR switched cervical segment.^{31,32} Moreover, other concerns such as higher probability of the incomplete removal of compressive hypertrophied bony or ligament structures or the exaggeration of the cervical mal-alignment when the ADR prosthesis is placed within the kyphotic segment might still exist.^{13,22}

But ADR has been repeatedly acclaimed from its capability to achieve the favorable clinical outcomes while permitting the physiological ROM maintenance without recurrence of pain or functional deficits after surgery in the studies dealing with the ADR applications over the pathologies (such as degenerative spondylosis, radiculopathy, or Modic change) other than CM.^{8,10,33,34} Long-term follow-up studies also showed comparable results with ACDF. A 48-month follow-up study of cervical radiculopathy (and/or myelopathy) has indicated the consistent, sustained

significantly superior outcomes for ADR when compared with ACDF.³⁵ The statistical superiority of ADR for overall success has persistently extended up to 7-year follow-up period that include biomechanical advantage of angular motion maintenance at the surgical level, less formation of the bridging bone, and lack of adjacent segment angulation increase either proximal or distal to the surgical level.³⁶ Overall, these results suggested that biomechanically mobile property of ADR might rather contribute to the better clinical and functional outcomes achievements, dispelling the suspicions such as feared worsening of symptomatic or neurological impairment from keeping its mobility instead of a firm stabilization at the compromised cervical segment even for prolonged period after surgery.

The results of current meta-analysis, the superiority of ADR over ACDF in terms of NDI and JOA, despite statistical insignificance, could be properly interpreted from the fact that, even though the mechanically thorough decompression of compressed spinal cord should be prioritized, the anatomical or physiological preservation of spinal column itself could partly contribute to the functional improvement for the CM subjects instead. Tian et al. have asserted that the limited neck ROM and its subsequent neck stiffness was main cause of disability for those undergoing ACDF and, instead, ADR could solve this problem.³⁷

Another concern might be the segmental kyphosis development or exaggeration after the ADR switch. The Bryan (Sofamor Danek, Memphis, TN, USA) prosthesis previously showed the less capacity to restore the regional lordosis due to the lack of anterior column support.^{38,39} However, more proper patient selection and technical modifications such as over-milling of the endplate's avoidance enabled the Bryan prosthesis to overcome this default.⁴⁰ Even ACDF also has the concerns of segmental kyphotic change, further leading to worse clinical outcomes, associated with subsidence of intervertebral cages.⁴¹

The patients in selected studies were mostly affected with single or 2 levels of CM and probably an extensive decompression was not mandated. This might have attributed to an excuse that the inadequate decompression from the incomplete removal of the neighboring structures compressing the spinal cord, one of the concerns after ADR for CM, was not brought up as a serious issue during this analysis. The multiple segment involvements caused more different pattern of biomechanical instability or substantial alignment change than single segment pathology.^{42,43} This analysis has supported that the ADR might be beneficial during the surgical management of few segments' involved CM, whereas its clinical implication on the multilevel CM that is often fraught with extensive ligament hypertrophy is

still yet to be supported. It could be assumed that, while the multiple segments arthroplasty might be inferior in terms of inadequacy of decompression or insufficient removal, multilevel fusion surgery might be vulnerable to a significant destabilization after the extensive decompression or might prompt a more severe adjacent disc degenerative process.^{7,44,45} Meanwhile, even in patients with nonlordotic alignment but without major kyphotic deformity, ADR had the potential to generate and maintain lordosis and improve patient-reported outcome measures in the short-term post-surgical results.⁴⁶

This study has several limitations. Firstly, the number of selected studies was small and available data was not sufficient for conducting meta-analysis in more extensive categories for a longer outcome period over 3 years. The primary distinction of this research from the existing literature might be the stepped follow-up points analysis. However, the inadequate number of studies at each follow-up point severely compromises the meta-analysis quality and caused serious statistical heterogeneity in some parameters. Secondly, there were differences in methodology across the studies, which might produce clinical heterogeneity. Thirdly, the restricted number of studies for each parameter was not enough to conduct subgroup analysis when there was significant heterogeneity. Fourthly, CIs in some categories were too widely ranged for achieving precision or accuracy. All these aspects lowered evidence level to low, consequently weaken the strength of meta-analysis. Further study with larger number of relevant articles in the future would be needed to provide the meta-analysis that would be more statistically powerful.

CONCLUSION

ADR was superior or at least, not inferior to ACDF in terms of functional recovery. However, its application to the CM patients is merely empowered with weak strength due to low level of evidence.

NOTES

Supplementary Material: Supplementary material can be found via <https://doi.org/10.14245/ns.2346498.249>.

Conflict of Interest: The authors have nothing to disclose.

Funding/Support: This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contribution: Conceptualization: JHL; Formal Anal-

ysis: YJL; Investigation: JHL; Methodology: MCC; Project Administration: JHL; Writing – Original Draft: JHL; Writing – Review & Editing: JHL.

ORCID

Jung Hwan Lee: 0000-0003-2680-6953

Youn Joo Lee: 0000-0002-5818-9542

Min Cheol Chang: 0000-0002-7629-7213

Jun Ho Lee: 0000-0001-6246-8053

REFERENCES

1. Lee JH, Lee SH, Seo IS. The characteristics of gait disturbance and its relationship with posterior tibial somatosensory evoked potentials in patients with cervical myelopathy. *Spine (Phila Pa 1976)* 2011;36:E524-30.
2. Chang CJ, Liu YF, Hsiao YM, et al. Comparison of anterior cervical discectomy and fusion versus artificial disc replacement for cervical spondylotic myelopathy: a meta-analysis. *J Neurosurg Spine* 2022;22:1-10.
3. Chen KT, Kim JS, Huang AP, et al. Current indications for spinal endoscopic surgery and potential for future expansion. *Neurospine* 2023;20:33-42.
4. Rao RD, Gourab K, David KS. Operative treatment of cervical spondylotic myelopathy. *J Bone Joint Surg Am* 2006;88:1619-40.
5. Yerneni K, Burke JF, Chunduru P, et al. Safety of outpatient anterior cervical discectomy and fusion: a systematic review and meta-analysis. *Neurosurgery* 2020;86:30-45.
6. Yoshii T, Egawa S, Chikuda H, et al. Comparison of anterior decompression with fusion and posterior decompression with fusion for cervical spondylotic myelopathy—a systematic review and meta-analysis. *J Orthop Sci* 2020;25:938-45.
7. Laratta JL, Shillingford JN, Saifi C, et al. Cervical disc arthroplasty: a comprehensive review of single-level, multilevel, and hybrid procedures. *Global Spine J* 2018;8:78-83.
8. Rožanković M, Marasanov SM, Vukić M. Cervical disk replacement with discover versus fusion in a single-level cervical disk disease: a prospective single-center randomized trial with a minimum 2-year follow-up. *Clin Spine Surg* 2017;30:E515-22.
9. Latka D, Kozłowska K, Miekisiak G, et al. Safety and efficacy of cervical disc arthroplasty in preventing the adjacent segment disease: a meta-analysis of mid- to long-term outcomes in prospective, randomized, controlled multicenter studies. *Ther Clin Risk Manag* 2019;15:531-9.

10. Hu X, Liu H, Wang B, et al. Cervical disc arthroplasty versus anterior cervical discectomy and fusion for the treatment of single-level disc degenerative disease with preoperative reversible kyphosis. *Clin Neurol Neurosurg* 2021;202:106493.
11. Ostrov PB, Reddy AK, Ryoo JS, et al. Anterior cervical discectomy and fusion versus cervical disc arthroplasty: a comparison of national trends and outcomes. *World Neurosurg* 2022;160:e96-110.
12. Liu ZQ, Hsieh CT, Huang CT, et al. Combining laminoplasty with artificial disc replacement for the treatment of cervical spondylotic myelopathy with congenital cervical stenosis. *Int J Spine Surg* 2023;30:8475.
13. Sekhon LH. Cervical arthroplasty in the management of spondylotic myelopathy. *J Spinal Disord Tech* 2003;16:307-13.
14. Brozek JL, Akl EA, Alonso-Coello P, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines. Part 1 of 3. An overview of the GRADE approach and grading quality of evidence about interventions. *Allergy* 2009;64:669-77.
15. Guyatt GH, Oxman AD, Kunz R, et al. What is "quality of evidence" and why is it important to clinicians? *BMJ* 2008; 336:995-8.
16. Faggion CM Jr. Evaluating the risk of bias of a study. *J Evid Based Dent Pract* 2015;15:164-70.
17. Kim SY, Park JE, Lee YJ, et al. Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *J Clin Epidemiol* 2013;66: 408-14.
18. Guyatt G, Gutterman D, Baumann MH, et al. Grading strength of recommendations and quality of evidence in clinical guidelines: report from an American college of chest physicians task force. *Chest* 2006;129:174-81.
19. Cheng L, Nie L, Li M, et al. Superiority of the Bryan disc prosthesis for cervical myelopathy: a randomized study with 3-year follow up. *Clin Orthop Relat Res* 2011;469:3408-14.
20. Ding C, Hong Y, Liu H, et al. Comparison of cervical disc arthroplasty with anterior cervical discectomy and fusion for the treatment of cervical spondylotic myelopathy. *Acta Orthop Belg* 2013;79:338-46.
21. Gornet MF, McConnell JR, Riew KD, et al. Treatment of cervical myelopathy: long-term outcomes of arthroplasty for myelopathy versus radiculopathy, and arthroplasty versus arthrodesis for myelopathy. *Clin Spine Surg* 2018;31: 420-7.
22. Riew KD, Buchowski JM, Sasso R, et al. Cervical disc arthroplasty compared with arthrodesis for the treatment of myelopathy. *J Bone Joint Surg Am* 2008;90:2354-64.
23. Shi S, Zheng S, Li XF, et al. Comparison of 2 zero-profile implants in the treatment of single-level cervical spondylotic myelopathy: a preliminary clinical study of cervical disc arthroplasty versus fusion. *PLoS One* 2016;11:e0159761.
24. Chen X, Shi L, Yu X, et al. Comparative study of artificial cervical disc replacement and anterior cervical discectomy/fusion in the treatment of cervical spondylotic myelopathy. *Int J Clin Exp Med* 2019;12:10597-604.
25. Bogduk N, Mercer S. Biomechanics of the cervical spine. I: normal kinematics. *Clin Biomech* 2000;15:633-48.
26. Hill P, Vaishnav A, Kushwaha B, et al. Comparison of Inpatient and outpatient preoperative factors and postoperative outcomes in 2-level cervical disc arthroplasty. *Neurospine* 2018;15:376-82.
27. Gao Y, Liu M, Li T, et al. A meta-analysis comparing the results of cervical disc arthroplasty with anterior cervical discectomy and fusion (ACDF) for the treatment of symptomatic cervical disc disease. *J Bone Joint Surg Am* 2013;95:555-61.
28. Lee JH, Lee JH. The discrepant clinical outcome predictions according to the differentiated centre of rotation shift after multilevel cervical total disc replacement. *Br J Neurosurg* 2020 Dec 28;1-6. doi: 10.1080/02688697.2020.1866162. [Epub].
29. Lee JH, Lee JH. The feasibility of optimal surgical result prediction according to the center of rotation shift after multilevel cervical total disc replacement. *Asian Spine J* 2020;14: 445-52.
30. Joaquim AF, Lee NJ, Riew KD. Revision surgeries at the index level after cervical disc arthroplasty - a systematic review. *Neurospine* 2021;18:34-44.
31. Nelson SY, Clark DM, Hoyt BW, et al. Cervical disk arthroplasty is an acceptable treatment option for cervical myelopathy. *Clin Spine Surg* 2022;35:95-6.
32. Baptiste DC, Fehlings MG. Pathophysiology of cervical myelopathy. *Spine J* 2006;6(6 Suppl):190S-197S.
33. Li J, Zhang D, Ding W, et al. Comparison of clinical outcomes after anterior cervical discectomy and fusion versus cervical total disk replacement in patients with modic-2 changes on MRI. *Clin Spine Surg* 2017;30:E1088-92.
34. Sakai K, Yoshii T, Arai Y, et al. Early experiences of one-level total disc replacement (Prestige LP) in Japan: a comparison of short-term outcomes with anterior cervical discectomy with fusion. *Spine Surg Relat Res* 2022;6:581-8.

35. Sasso RC, Anderson PA, Riew KD, et al. Results of cervical arthroplasty compared with anterior discectomy and fusion: Four-year clinical outcomes in a prospective, randomized controlled trial. *J Bone Joint Surg Am* 2011;93:1684-92.
36. Gornet MF, Burkus JK, Shaffrey ME, et al. Cervical disc arthroplasty with Prestige LP disc versus anterior cervical discectomy and fusion: seven-year outcomes. *Int J Spine Surg* 2016;10:24.
37. Tian W, Yan K, Han X, et al. Comparison of the clinical and radiographic results between cervical artificial disk replacement and anterior cervical fusion: a 6-year prospective non-randomized comparative study. *Clin Spine Surg* 2017;30:E578-86.
38. Pickett GE, Mitsis DK, Sekhon LH, et al. Effects of a cervical disc prosthesis on segmental and cervical spine alignment. *Neurosurg Focus* 2004;17:E5.
39. Johnson JP, Laurysen C, Cambron HO, et al. Sagittal alignment and the Bryan cervical artificial disc. *Neurosurg Focus* 2004;17:E14.
40. Yanbin Z, Yu S, Zhongqiang C, et al. Sagittal alignment comparison of Bryan disc arthroplasty with ProDisc-C arthroplasty: a prospective, randomized controlled clinical trial. *J Spinal Disord Tech* 2011;24:381-5.
41. Duey AH, Gonzalez C, Geng EA, et al. The effect of subsidence on segmental and global lordosis at long-term follow-up after anterior cervical discectomy and fusion. *Neurospine* 2022;19:927-34.
42. Lee JH, Lee SH. Clinical and radiological characteristics of lumbosacral lateral disc herniation in comparison with those of medial disc herniation. *Medicine (Baltimore)* 2016;95:e2733.
43. Ding SF, Zhang ZY, Jiang ZJ, et al. Biomechanical evaluation of cervical spine instability after multiple level laminectomy. *Chin Med J (Engl)* 1991;104:626-33.
44. Jiang L, Tan M, Yang F, et al. Comparisons of safety and clinical outcomes between multiple-level and single-level cervical disk replacement for cervical spondylosis: a systematic review and meta-analysis. *Clin Spine Surg* 2016;29:419-26.
45. Samuel AM, Moore HG, Vaishnav AS, et al. Effect of myelopathy on early clinical improvement after cervical disc replacement: a study of a local patient cohort and a large national cohort. *Neurospine* 2019;16:563-73.
46. Mok JK, Vaishnav AS, Chaudhary C, et al. Impact of non-lordotic sagittal alignment on short-term outcomes of cervical disc replacement. *Neurospine*. 2020;16:588-602.