Anterior Cervical Discectomy and Fusion versus Cervical Disc Arthroplasty for Cervical Degenerative Disc Disease: Alarming Rates of Bias

Joshua Davood[1], Henry Avetisian[1], Daniel Rusu[1], Marco Napolitano[1], Joshua Khorsandi[1,2], Apurva Prasad[1], Joseph Borna[1], Joseph Noorvash[1], Nicole Kahen[1], Aidan Lindgren1, William J. Karakash1, Ram K. Alluri MD1

[1]Department of Orthopedic Surgery, University of Southern California (USC) Keck School of Medicine, Los Angeles, USA

[2]Kirk Kerkorian School of Medicine at UNLV, Las Vegas, USA

davood@usc.edu

Disclosures: **J. Davood**: None. **H. Avetisian**: None. D**. Rusu**: None. **M. Napolitano**: None. **J. Khorsandi**: None. **A. Prasad**: None. **J. Borna**: None. **J. Noorvash**: None. **N. Kahen**: None. **A. Lindgren**: None. **W.J. Karakash**: None. **R.K. Alluri**: 2; eCential Robotics. 3B; Medtronic, Orthofix, Globus. 4; HIA Technologies, NeoOnc, Globus. 5; Max BioPharma, ATEC.

INTRODUCTION: Cervical degenerative disc disease (CDDD) is commonly treated with anterior cervical discectomy and fusion (ACDF) or cervical disc arthroplasty (CDA), but their comparative effectiveness remains debated. Systematic reviews and meta-analyses are crucial for guiding clinical decision-making, yet spin bias—selective framing that overstates benefits or downplays limitations—remains prevalent in orthopaedics and can mislead clinicians. We aimed to quantify the prevalence and types of spin and to appraise methodological quality among systematic reviews and meta-analyses comparing CDA and ACDF. We hypothesized that spin bias was present in at least 40% of systematic reviews and meta-analyses comparing ACDF and CDA.

METHODS: A systematic search of PubMed, Web of Science, and Embase was conducted in September 2024 to identify systematic reviews and meta-analyses. Inclusion required: meta-analysis or systematic review; CDDD; direct single- or multi-level ACDF vs CDA comparison of treatment; available in English. Two independent reviewers screened titles/abstracts and full texts with third-reviewer adjudication, then assessed the included articles for the nine most spin categories defined by Yavchitz et al.[1] and AMSTAR-2 (a methodological quality assessment tool). Journal impact factor, publication year, yearly/total citations, and funding were extracted for each article. Associations between AMSTAR-2 rating, yearly citations, and journal impact factor and the number of types of spin present or presence/absence of spin were assessed using Poisson regression modeling and binary logistic regression modeling, respectively.

RESULTS: Screening yielded 351 articles; 77 articles (2010–2024) met inclusion criteria. Spin was present in 49/77 (64%) articles; number of spin types per article ranged 0–4. The most frequently observed spin type was Type 5, found in 30/77 (39%) of the articles. Type 3 was found in 20/77 (26%), Type 6 in 12/77 (16%), Types 9 and 4 in 7/77 (9%), Type 1 in 5/77 (6%), Type 7 in 4/77 (5%), and Type 8 in 1/77 (1%) (Table 1). No articles demonstrated spin type 2. All spin-positive articles favored CDA. AMSTAR-2 assessment rated 1/77 (1%) articles as high, 0 as moderate, 12/77 (16% as low, and 64/77 (83%) as critically low. Journal impact factors ranged from 0.3-10.1 (mean 2.53); 28/77 (36%) studies reported receiving funding; yearly citations ranged from 0–18.71 (mean 4.65) (Table 2). Our binary logistic regression model found AMSTAR-2 rating to significantly predict presence of spin (OR: 0.182 [0.048-0.689], p=0.012) while yearly citations and impact factor did not (both p>0.05). Lastly, our Poisson regression model found yearly citations and impact factor did not significantly predict amount of spin (both p>0.05) while AMSTAR-2 rating fell just short of significant correlation (IRR: 0.517 [0.261-1.023], p=0.058).

DISCUSSION: Spin bias was strongly apparent in CDA vs ACDF systematic reviews and meta-analyses. Type 5 was the most prevalent type of spin found, indicating that abstract and article conclusions often outpaced the certainty of the underlying evidence. AMSTAR-2 ratings were generally poor due to the strictness of criteria, but emphasize the need for improving and standardizing methodology in orthopaedic research. Lower AMSTAR-2 ratings were associated with the presence of spin, whereas journal impact factor and citation counts were not—suggesting editorial prestige and popularity do not safeguard against biased framing. Spin-positive articles uniformly favored CDA, suggesting there may be external influences such as industry funding driving bias. Limitations include subjectivity in spin classification and restriction to English-language reviews.

SIGNIFICANCE/CLINICAL RELEVANCE: Severe spin was common, and methodological quality was frequently low among CDA vs ACDF reviews. Clinicians should interpret abstract claims cautiously and appraise full texts for strong methodology and transparency in both reviews and included studies. Greater transparency and standardized reporting will reduce misinterpretation and better support decisions between CDA and ACDF.

REFERENCES: 1. Yavchitz A, Ravaud P, Altman DG, et al. A new classification of spin in systematic reviews and meta-analyses was developed and ranked according to the severity. *J Clin Epidemiol*. 2016;75:56-65.

Types of Spin	Articles with Spin: n (%)
1. The conclusion formulates recommendations for clinical practice not supported by the findings	5 (6.49%)
2. The title claims or suggests a beneficial effect of the experimental intervention not supported by the findings	0 (0%)
3. Selective reporting of or overemphasis on efficacy outcomes or analysis favoring the beneficial effect of the experimental intervention	20 (25.97%
4. The conclusion claims safety based on non-statistically significant results with a wide confidence interval	7 (9.09%)
5. The conclusion claims the beneficial effect of the experimental treatment despite a high risk of bias in primary studies	30 (38.96%)
6. Selective reporting of or overemphasis on harm outcomes or analysis favoring the safety of the experimental intervention	12 (15.58%)
7. The conclusion extrapolates the review findings to a different intervention	4 (5.20%)
8. Conclusion extrapolates the review's findings from a surrogate marker or a specific outcome to the global improvement of the disease	5 (6.49%)
9. Conclusion claims the beneficial effect of the experimental treatment despite reporting bias	1 (1.30%)

Table 1. Prevalence of the Nine Most Severe Types of Spin per Yavchitz et al.

	Spin (n = 49)	No Spin (n = 28)
AMSTAR-2 Rating		
High	0	1
Moderate	0	0
Low	4	8
Critically Low	45	19
Journal Impact Factor	2.53 ± 1.55	2.54 ± 1.01
Yearly Citations	4.80 ± 4.08	4.40 ± 2.75

Table 2. Characteristics of Studies With and Without Spin Present