The Statistical Fragility of Cervical Disc Arthroplasty versus Anterior Cervical Discectomy and Fusion for Degenerative Cervical Pathology: A Systematic Review and Meta-analysis

Carlos David Ortiz-Babilonia[1], Arjun Gupta, Amy Li Xu, Wesley Michael Durand[2], Micheal Raad, Amit Jain
[1]Orthopaedic Surgery, Johns Hopkins University, [2]Brown University

INTRODUCTION: Anterior cervical discectomy and fusion (ACDF) has long been considered the gold standard for treatment of degenerative cervical pathology (DCP) resulting in radiculopathy and/or myelopathy. Cervical disc arthroplasty (CDA) was later developed for this indication as an alternative to account for shortcomings associated with fusion, including decreased mobility and adjacent segment disease (ASD). Randomized controlled trials comparing these interventions for this indication have consistently shown that CDA is equivalent and may even be more effective than ACDF. Since RCTs are the highest level of original research available, these findings may heavily influence how spine surgeons decide to treat symptomatic DCP. However, noted lack of reproducibility of RCTs has brought special attention to how they determine significance, casting doubt on the almost ubiquitously utilized $p<0.05$ threshold for statistical significance. In this meta-analysis, we assessed the statistical robustness of RCTs comparing cervical disc arthroplasty to anterior cervical discectomy and fusion for the treatment of DCP by using the continuous fragility index (CFI) and quotient (CFQ).

METHODS:

RCTs evaluating outcomes after CDA versus ACDF were included. Dichotomous and continuous outcomes were collected from each study. Incidence of adjacent segment disease (ASD) was the only dichotomous outcome evaluated. Continuous outcomes included Neck Disability Index (NDI), overall Visual Analogue Scale (VAS), neck VAS, arm VAS, and modified Japanese Orthopaedic Association (mJOA) score. The dichotomous or continuous fragility index (FI or CFI) was determined by manipulating each outcome measure until reversal of significance (a=0.05) was achieved. The corresponding fragility quotients (FQ) were calculated by dividing the FI or CFI by the sample size.

RESULTS: Of 1,747 studies screened, 20 studies (13 dichotomous outcome events; 56 continuous events) were included. The median FI for dichotomous events (ASD occurrence) was 7, suggesting that altering the outcome of 7 patients would be required to reverse trial significance. For continuous events, altering the treatment of 13 patients would be required to reverse significance. The corresponding FQs were 0.043 and 0.123 for dichotomous and continuous events, respectively. By prosthesis type, RCTs reporting on ActivC®, Prestige®, and Discover® devices had the highest CFQ values and therefore the greatest statistical robustness, whereas data for PCM Cervical Disc® and Secure-C® implants were comparatively weaker.

DISCUSSION AND CONCLUSION: Our findings suggest that RCTs evaluating CDA versus ACDF for treatment of DCP are statistically robust. This study should provide reassurance to spine surgeons and patients that level 1 findings regarding CDA in comparison to ACDF can be trusted, and thus CDA may be at least as effective as ACDF for the treatment of DCP with radiculopathy and/or myelopathy. Given the importance of RCTs in clinical decision-making, fragility indices should be reported alongside p-values to indicate the strength of statistical findings.

Table 1. Statistical fragility of adjacent segment disease (ASD) outcomes in RCTs evaluating CDA versus ACDF

	Events	Patients	LostFU	FI	FQ
All trials	13	1859	753	7 (3–10)	0.043 (0.035–0.066)
Outcome					
ASD	4	638	25	10.5 (3.25–19)	0.05 (0.03–0.1)
Superior ASD	4	601	296	5 (3–22)	0.041 (0.035–0.116)
Inferior ASD	5	620	432	9 (3–9)	0.043 (0.038–0.066)
Prosthesis					
activC	2	132	10	2.5 (2.25–2.75)	0.038 (0.034–0.042)
Bryan	0	0			
Discover	1	114	21	4 (4–4)	0.035 (0.035–0.035)
Mobi-C	5	631	466	9 (3–10)	0.066 (0.038–0.078)
PCM Cervical Disc	0	0			
Prestige	0	0			
Secure-C	1	380	4	25 (25–25)	0.066 (0.066–0.066)
Original P value					
p < 0.05	5	555	181	3 (3–7)	0.038 (0.028–0.038)
p > 0.05	8	1304	572	9 (3.75–13.75)	0.056 (0.041–0.069)

Table 2. Statistical fragility of continuous outcomes in RCTs evaluating CDA versus ACDF

	Events	Patients	LostFU	CFI (median, IQR)	CFQ (median, IQR)
All trials	56	8324	1787	13 (6.75–21.25)	0.123 (0.065–0.182)
Outcome					
NDI	18	2708	575	13 (8–22)	0.135 (0.067–0.179)
VAS	6	410	13	9.5 (4–12.75)	0.182 (0.112–0.199)
Neck VAS	13	2352	579	14 (11–21)	0.11 (0.061–0.157)
Arm VAS	14	2473	612	16.5 (7.25–23.5)	0.112 (0.082–0.16)
mJOA	5	381	8	9 (3–13)	0.114 (0.063–0.182)
Prosthesis					
ActivC	4	264	20	13 (12.5–13.25)	0.197 (0.189–0.201)
Bryan	3	327	33	12 (9.5–15)	0.11 (0.087–0.138)
Discover	13	1476	252	13 (5–22)	0.125 (0.094–0.218)
Mobi-C	15	2730	639	20 (14.5–26)	0.112 (0.073–0.151)
PCM Cervical Disc	3	904	229	17 (16–19.5)	0.065 (0.058–0.065)
Prestige	6	1222	458	13 (4.25–22.5)	0.139 (0.066–0.234)
Secure-C	4	900	156	5.5 (4.25–7.5)	0.024 (0.019–0.033)
Original P value					
p < 0.05	8	1704	437	21.5 (18.25–23)	0.148 (0.083–0.205)
p > 0.05	46	6445	1334	12.5 (5.25–17.75)	0.117 (0.064–0.178)