Adjacent Segment Pathology After Treatment With Cervical Disc Arthroplasty or Anterior Cervical Discectomy and Fusion, Part 1: Radiographic Results at 7-Year Follow-Up

PIERCE D. NUNLEY, MD, EUBULUS J. KERR III, MD, DAVID A. CAVANAUGH, MD, PHILLIP ANDREW UTTER, MD, PETER G. CAMPBELL, MD, RISHI WADHWA, MD, KELLY A. FRANK, MS, KYLE E. MARSHALL, MS, MARCUS B. STONE, PHD
Spine Institute of Louisiana, Shreveport, Louisiana

ABSTRACT
Background: Adjacent segment pathology (ASP) remains a concern following treatment with cervical disc arthroplasty (CDA) and anterior cervical discectomy and fusion (ACDF). Radiographic ASP (RASP) is ASP identified on imaging, which may or may not include clinical symptoms. The risk factors for development of RASP and its clinical effects remain controversial. In part 1 of a 2-part publication we evaluate the incidence and predictors of RASP as well as determine whether any association exists between RASP and patient-reported outcomes (PROs).

Methods: Data were prospectively collected during a US Food and Drug Administration randomized, multicenter, investigational device exemption trial comparing CDA (Mobi-C; Zimmer Biomet, Westminster, CO) with ACDF. Multiple post hoc analyses were conducted on RASP as it related to demographics and patient outcomes. Kaplan-Meier estimates of time to Kellgren-Lawrence (K-L) grade 3/4 were calculated separately for all groups. Multivariate Cox proportional hazard models were used analyze whether RASP was associated with patient preoperative demographic characteristics and preoperative and postoperative radiographic characteristics. The association of RASP with PROs was analyzed using generalized estimating equations and matched, retrospective cohort analysis.

Results: The incidence of grade 3/4 RASP was lower for patients treated with CDA when initial treatment was at 1 level (27% vs 47%, \( P < .0001 \)) and at 2 levels (14% vs 49%, \( P < .0001 \)). Kaplan-Meier estimates indicated significantly lower probability of grade 3/4 RASP over time for patients receiving CDA (\( P < .001 \)). Treatment with ACDF, treatment of 1 level, higher age, body mass index, higher preoperative physical components score, and a lower Cobb angle were associated with elevated risk of grade 3/4 RASP. CDA was shown to be more effective than ACDF (64.4%; 95% CI = 50.9, 74.2; \( P < .0001 \)) at preventing RASP.

Conclusions: The incidence and risk of RASP is decreased when patients are treated with CDA compared with ACDF. Although the mechanism of CDA that generates this protective effect is not understood, PROs remain unaffected through 7 years despite changes in RASP.

Cervical Spine

Keywords: adjacent segment pathology, cervical disc arthroplasty, anterior cervical discectomy and fusion

INTRODUCTION
Adjacent segment pathology (ASP) has been recognized as an important sequelae following both anterior cervical discectomy and fusion (ACDF) and cervical disc arthroplasty (CDA) surgeries. ACDF is known to increase the mechanical stress and displacement of motion on the segments adjacent to the fused level(s), whereas the motion-sparing design of CDA reduces this stress and displacement.\(^1,2\) The reduction of stress and displacement by CDA at the adjacent level has been hypothesized\(^2-4\) to reduce the incidence of ASP. As a result, a number of studies\(^5-7\) have made efforts to compare CDA and ACDF rates of ASP.

ASP has been analyzed and reported using radiographic measures and clinical outcomes (CASP). Radiographic adjacent segment pathology (RASP), the focus of this analysis, represents the development of new, radiographically identified, degenerative changes adjacent to the treated level. CASP is the development of clinical symptoms at the level adjacent to the previously treated level. CASP will be analyzed in part 2 of this publication.

RASP has been defined using a variety of measurements and scales. Using magnetic resonance...
imaging (MRI), computed tomography (CT), static x-rays, and dynamic x-rays, authors\textsuperscript{8–11} have developed grading scales for measuring RASP in the cervical spine, yet no standardized method exists. Some measures of RASP reported within the literature\textsuperscript{12–15} include range of motion (ROM) at the index and adjacent levels as well as the entire cervical spine, plate-to-disc distances, disc bulge impingements, osteophyte formation, heterotopic ossification formation, and narrowing of the disc space.

Multiple studies\textsuperscript{6,13,16–26} have reported that treatment with ACDF resulted in higher incidence rates of RASP than of CDA, although other studies\textsuperscript{5,27} found no difference. Reported RASP rates ranged from 7\% to 92\% of patients following ACDF surgery and from 3.8\% to 71\% of patients following CDA.\textsuperscript{28–30} The pivotal Hilibrand et al\textsuperscript{2} paper reported RASP via MRI scans but only analyzed the inverse relationship of patients who already had radicular or myelopathic symptoms (CASP) after ACDF surgery. Debate continues to surround the true incidence and relevance of RASP after CDA and ACDF.

Studies\textsuperscript{12,28,31,32} have demonstrated independently that the development of RASP after CDA or ACDF may be influenced by myriad factors. Age has been a particularly contested factor among ASP studies, with a recent meta-analysis determining that younger age contributed to more RASP after ACDF. However, Hilibrand et al\textsuperscript{33} also found a significant correlation between patient age and RASP, with older patients tending to develop grade III or IV degenerative changes at adjacent segments (average age of 64.5 years). Interesting findings in the literature have also shown that surgery of only 1 level with either CDA or ACDF places a patient at higher risk for RASP than a multilevel surgery.\textsuperscript{28} We were not able to find any significant associations between RASP and patient-reported outcomes (PROs) within the literature. An overwhelming majority of studies either did not report these results\textsuperscript{14,27,32,34,35} or, when a connection between these variables was analyzed, did not find a correlation between PROs (eg, visual analog pain scores and neck disability index) and RASP.\textsuperscript{31}

In part 1 of this 2-part publication, we evaluate the incidence of RASP at 7 years in patients treated with CDA and ACDF at 1 and 2 levels, evaluate predictors of RASP, and analyze whether an association exists between RASP and PROs. Part 2 of this publication will focus on CASP.

**METHODS**

All data were prospectively collected during a US Food and Drug Administration (FDA) randomized, multicenter, investigational device exemption (IDE) trial comparing CDA (Mobi-C; Zimmer Biomet, Westminster, CO) with ACDF. Patients were diagnosed with symptomatic degenerative disc disease at 1 or 2 contiguous levels from C3-7. Enrollment in the 1-level arm included 164 patients treated with CDA and 81 with ACDF, whereas the 2-level arm included 225 treated with CDA and 105 with ACDF. Details and overall results of the trial have been reported previously.\textsuperscript{6,22–26}

**Study Design**

Part 1 of this multiphase analysis was a post hoc study of radiographic and clinical data over 7 years as it relates to RASP.

Demographics were collected including age, race, gender, height, weight, and body mass index (BMI). PROs included a neck disability index (NDI), visual analog scale (VAS) arm and neck, and the 12-item Short-Form Health Survey mental and physical components (SF-12 MCS and SF-12 PCS).

All available radiographs were analyzed preoperatively and postoperatively at years 1 to 5 and year 7. Radiographic evaluations of RASP were performed by independent radiologists (Medical Metrics Inc, Houston, TX). RASP was evaluated at both the inferior and superior adjacent level according to the K-L scale (Table 1) as modified for cervical spine.\textsuperscript{9,36}

**Statistical Methods**

**Radiographic Adjacent Segment Pathology**

For each subject and time point, RASP was defined as the maximum K-L grade of either the inferior or superior adjacent level. Kaplan-Meier estimates of time to K-L RASP grade 3/4 were calculated.

<table>
<thead>
<tr>
<th>K-L Grading Scale:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 0: No degeneration</td>
</tr>
<tr>
<td>Grade 1: Minimal anterior osteophytosis</td>
</tr>
<tr>
<td>Grade 2: Definite anterior osteophytosis, possible narrowing of disc space</td>
</tr>
<tr>
<td>Grade 3: Moderate narrowing of disc space, sclerosis, osteophytosis</td>
</tr>
<tr>
<td>Grade 4: Severe narrowing of disc space, sclerosis, large osteophytes</td>
</tr>
</tbody>
</table>

*International Journal of Spine Surgery, Vol. 14, No. 3 270*
separately for the ACDF and CDA groups and for the 1-level and 2-level cohorts. The Kaplan-Meier estimates represent the probability of having grade 3/4 RASP at each postoperative time point. The Kaplan-Meier estimates were compared using the log-rank test. Poisson incidence rates of grade 3/4 RASP and corresponding confidence intervals (CIs) were included on each Kaplan-Meier plot and presented as the number of new cases of grade 3/4 RASP as a percentage per patient-year (%/pt-yr).

Multivariate Cox proportional hazard models were used to determine whether time to grade 3/4 RASP was associated with patient preoperative demographic characteristics and preoperative and postoperative radiographic characteristics. Preoperative covariates included treatment group, levels treated, age, sex, race, BMI, NDI score, and SF-12 MCS and PCS scores. Postoperative covariates included C2-7 Cobb angle at 6 weeks, mean flexion-extension ROM of the index level(s) at 3 months, and mean functional spinal unit (FSU) height of the index level(s) at 6 weeks postoperation. Interactions between treatment and the other covariates were assessed in a serial fashion with bivariate models to determine whether covariates modified the effect of treatment. Estimates of CDA efficacy in preventing RASP were defined as 1 minus the hazard ratio for CDA versus ACDF and are presented as percentages. An efficacy of ≤ 0% indicates no benefit of CDA in preventing RASP compared with ACDF, whereas an efficacy near 100% indicates great benefit.37

Association of RASP With PROs: Generalized Estimating Equation
To determine whether RASP had a significant association with PROs of pain and function, a series of weighted generalized estimating equation models were constructed. Inverse probability weighting was used to adjust for the effect of missing data due to missed visits, patient drop out, or secondary surgery. Models were defined using NDI, VAS neck pain, SF-12 PCS, or SF-12 MCS score as the dependent variable, and follow-up time and RASP as the independent variables. RASP was modeled on both the 5-point K-L scale and on the binary scale defined above. A term for the reciprocal of follow-up month was included to account for the improvement in PRO scores following surgery. Interactions between RASP and time were included to determine whether the effects of RASP were dependent on postoperative time. In addition, a series of multivariate models were fit as sensitivity analyses to determine whether other patient characteristics had an impact on the grade of RASP.

Association of RASP With PROs: Matched, Retrospective Cohort Analysis
As a second approach to determine the association between PRO scores and RASP, we conducted a matched, retrospective cohort analysis. RASP acquisition was fixed at a time point to control for the time spent with RASP. Cases were selected as all participants with grade 3/4 RASP by month 24. Cases were matched to controls in 1:3 ratio using a nearest-neighbor method, matching on sex, levels treated, treatment, age, and baseline NDI score. Binary RASP models were fit using the same methods and weights described above. In addition, controls were censored at the time they reported grade 3/4 RASP or underwent a secondary surgery. P values less than .05 are considered statistically significant.

RESULTS
Radiographic Adjacent Segment Pathology
At 7 years, the incidence of grade 3/4 RASP for the 1-level patients was 27% for CDA and 47% for ACDF (P < .0001); for the 2-level patients it was 14% for CDA and 49% for ACDF (P < .0001; Fisher exact test).

Kaplan-Meier estimates (Figures 1 and 2) showed significantly higher probability of grade 3/4 RASP over time for patients treated with ACDF compared
with those treated with CDA for both the 1- and 2-level cohorts \((P < .001)\).

For the 1-level cohort, the annual incidences of grade 3/4 RASP were 5.1%/pt-yr (CI = 3.7, 7.0) and 11.1%/pt-yr (CI = 7.7, 15.5) for the CDA and ACDF groups, respectively. For the 2-level cohort the annual incidences of grade 3/4 RASP were 2.7%/pt-yr (CI = 1.9, 3.9) and 9.9%/pt-yr (CI = 7.1, 13.3) for the CDA and ACDF groups, respectively.

A multivariate Cox model showed that treatment with ACDF, treatment of only 1 level, higher age, higher BMI, higher preoperative PCS score, and a lower Cobb angle were associated with elevated risk of grade 3/4 RASP (Table 2). The overall efficacy of CDA in preventing RASP compared with ACDF was 64.4% (95% CI = 50.9, 74.2; \(P < .0001\)). There were no significant effect modifications of treatment by the other covariates (Table 3).

**DISCUSSION**

Symptomatic adjacent segment pathology is generally accepted as a sequelae of cervical spine surgery with ACDF and CDA; however, much debate still surrounds the true incidence and relevance of RASP.

**RASP Rates**

Although variances in methods present a challenge, the rates of RASP reported here for CDA and ACDF (respectively, 27% versus 47% at 1 level and 14% versus 49% at 2 level) appear reasonable when compared with the current literature.

The rates of RASP reported in the literature have been collected using various radiographic methods including plain x-ray, flexion-extension x-ray, CT, and MRI. There is no commonly accepted method of RASP measurement. There are some established examples including K-L, Miyazaki et al, Matsumoto et al, and Thompson, but often RASP measurements are defined per study as combina-
tions of osteophyte formation or changes, foraminal narrowing, new or increased calcification of the anterior longitudinal ligament, nucleus color, structure, disk bulge, and disc height.29,30,40,41

Several recently published meta-analyses pooled results to define RASP rates across cervical studies. Due to the variety of collection and measurement techniques for RASP, these results must be cited and interpreted with caution. In a 2013 review, Carrier et al.29 reported an average RASP rate of 27.3% (range, 16.1%–71.4%) for 553 patients with 1- or 2-level ACDF. A 2016 review30 reported pooled RASP for CDA, with an average of 8.3% (95% CI = 3.8%, 12.7%). An additional review28 was published in 2018, with Hashimoto reporting an average RASP rate of 32.8% (range, 7%–92%) across 34,716 patients who had undergone cervical fusion.

Reported results from individual FDA IDE studies do not consistently include RASP. Although the data collected for these trials were robust, the studies were not designed to analyze RASP. Analyses of RASP from these studies were commonly post hoc, as is the case with our analyses. The IDE studies with reported RASP rates include Kineflex-C and PCM. The Kineflex-C study included RASP rates measured by disc space narrowing, osteophytes, and endplate sclerosis through 60 months. RASP was in favor of CDA at the superior (17.1% vs 32.2%, \( P < .01 \)) and inferior levels (24.5% vs 28.9%, not significant).42 RASP from the PCM IDE was reported as any worsening from baseline in disc height loss, presence or size of

<p>| Table 3. Estimated incidence of grade 3/4 RASP and efficacy of CDA in preventing grade 3/4 RASP from months 0–84, by participant characteristics.a |
|----------------|----------------|----------------|----------------|</p>
<table>
<thead>
<tr>
<th>Variablea</th>
<th>ACDF n RASP 3/4</th>
<th>CDA n RASP 3/4</th>
<th>CDA Efficacy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Person-years</td>
<td>Rateb</td>
<td>% (95% CI)</td>
</tr>
<tr>
<td>Overall</td>
<td>186 76</td>
<td>731.4</td>
<td>10.4</td>
</tr>
<tr>
<td>Levels treated</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>81 34</td>
<td>306.1</td>
<td>11.1</td>
</tr>
<tr>
<td>2</td>
<td>105 42</td>
<td>425.3</td>
<td>9.9</td>
</tr>
<tr>
<td>Age, y</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21–34</td>
<td>12 4</td>
<td>45.5</td>
<td>8.8</td>
</tr>
<tr>
<td>35–49</td>
<td>123 45</td>
<td>491.0</td>
<td>9.2</td>
</tr>
<tr>
<td>50–67</td>
<td>51 27</td>
<td>194.9</td>
<td>13.9</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>105 41</td>
<td>412.4</td>
<td>9.9</td>
</tr>
<tr>
<td>Male</td>
<td>81 35</td>
<td>319.0</td>
<td>11.0</td>
</tr>
<tr>
<td>Racec</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>18 11</td>
<td>59.8</td>
<td>18.4</td>
</tr>
<tr>
<td>White</td>
<td>168 65</td>
<td>671.6</td>
<td>9.7</td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 30</td>
<td>134 55</td>
<td>523.6</td>
<td>10.5</td>
</tr>
<tr>
<td>≥ 30</td>
<td>52 21</td>
<td>207.9</td>
<td>10.1</td>
</tr>
<tr>
<td>NDI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 50</td>
<td>69 33</td>
<td>266.1</td>
<td>12.4</td>
</tr>
<tr>
<td>≥ 50</td>
<td>117 43</td>
<td>465.3</td>
<td>9.2</td>
</tr>
<tr>
<td>SF-12 PCS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 35</td>
<td>120 47</td>
<td>487.7</td>
<td>9.6</td>
</tr>
<tr>
<td>≥ 35</td>
<td>66 29</td>
<td>243.7</td>
<td>11.0</td>
</tr>
<tr>
<td>SF-12 MCS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 45</td>
<td>107 35</td>
<td>413.8</td>
<td>8.5</td>
</tr>
<tr>
<td>≥ 45</td>
<td>79 41</td>
<td>317.6</td>
<td>12.9</td>
</tr>
<tr>
<td>C2–7 Cobb angle</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angle &lt; −5</td>
<td>24 9</td>
<td>95.2</td>
<td>9.5</td>
</tr>
<tr>
<td>−5 ≤ Angle ≤ 5</td>
<td>37 16</td>
<td>138.2</td>
<td>11.6</td>
</tr>
<tr>
<td>Angle &gt; 5</td>
<td>125 51</td>
<td>498.0</td>
<td>10.2</td>
</tr>
<tr>
<td>ROM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 2º</td>
<td>129 56</td>
<td>529.3</td>
<td>10.6</td>
</tr>
<tr>
<td>≥ 2º</td>
<td>57 20</td>
<td>202.1</td>
<td>9.9</td>
</tr>
<tr>
<td>FSU height</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 30 mm</td>
<td>111 42</td>
<td>415.8</td>
<td>10.1</td>
</tr>
<tr>
<td>≥ 30 mm</td>
<td>75 34</td>
<td>315.6</td>
<td>10.8</td>
</tr>
</tbody>
</table>

Abbreviations: ACDF, anterior cervical discectomy and fusion; BMI, body mass index; CDA, cervical disc arthroplasty; FSU, functional spinal unit; MCS, mental components score; NDI, neck disability index; PCS, physical components score; RASP, radiographic adjacent segment pathology; ROM, range of motion; SF-12, Short-Form Health Survey with 12 items.

aDemographic and patient-reported outcomes were taken from preoperative assessments. Radiographic measures are for the first available postoperative assessment.
bRate = \( \frac{100}{3} \times (\text{No. with RASP} / \text{No. Person-yrs}) \).
cFor race/ethnicity, Asian, Black, Hawaiian/Pacific Islander, Hispanic, Native American/Alaska Native, and multiracial individuals are grouped together as “Other” due to small sample size.

osteophytes, and endplate sclerosis, with significantly different superior adjacent RASP rates of 33.1% for CDA and 50.9% for ACDF \( (P = .006) \), and similar inferior RASP rates of 49.2% for CDA and 51.7% for ACDF \( (P = .78) \).\(^{44}\) Several IDE studies did not analyze RASP but did compare the adjacent segment ROM between CDA and ACDF.\(^{5,45–55}\) The results ranged from no difference between Prestige ST and ACDF at 7 years,\(^{5}\) to increased ROM for ACDF when compared with 1-level Prestige LP at 7 years.\(^{56}\) Overall, the hypothesized correlation between increased ROM at the adjacent level and incidence of RASP has not been proven in an IDE study or otherwise.\(^{14,20,27,51,57}\)

RASP Predictors

Predictors of grade 3/4 RASP were identified as treatment with ACDF, treatment at only 1 level, higher age, higher preoperative SF-12 PCS, and a lower Cobb angle. Treatment with ACDF is hypothesized to increase the risk of RASP in part by increased motion and stresses at the adjacent level. As discussed previously, increased ROM remains uncorrelated to RASP. The increased risk of RASP with treatment at only 1 level is supported in the literature. We believe patients with multilevel disease may initially be symptomatic at only 1 level and be treated conservatively. Our analysis indicated increased risk of RASP with higher age,\(^{2}\) but the literature is divided: Some studies support our findings, whereas others find correlation to younger age.\(^{28}\) On the basis of the data reported by Boden et al,\(^{58}\) we hypothesize that older patients have a higher likelihood of nonsymptomatic RASP as part of the natural course of aging.

RASP Relevance

The clinical relevance of RASP remains a topic that is heavily debated. We did not observe a correlation between a higher incidence of RASP with poor PROs. Comparison to the literature is difficult, given that most studies report RASP independent of PROs and CASP.\(^{29–31,59–61}\) The correlation of RASP, in this data set, with CASP, defined as adjacent-level secondary surgery, is included in part 2 of this paper.

Boden et al tested the validity of RASP by comparing radiographic evaluations of symptomatic and asymptomatic patients. The MRIs of 37 patients with a confirmed symptomatic abnormality and 63 asymptomatic patients were reviewed by 3 neuroradiologists. A major abnormality was noted in 19% of the asymptomatic patients. Asymptomatic patients over 40 years were more likely to be interpreted, on the basis of radiographs, as having a major abnormality than were those under 40 (28% and 14%, respectively). Disc degeneration occurring at 1 level or more was also more prevalent in patients over 40 years, with 60% being interpreted as having an abnormality, whereas this interpretation applied to only 25% of patient under 40 years of age. The authors cautioned against operative decisions based solely on radiographic findings, because changes could be part of the normal aging process.\(^{58}\)

This analysis contains limitations because it is a post hoc analysis of a prospective study, not a prospectively planned analysis of RASP. Radiographic variables were fixed at early time points to determine whether early radiographic measurements could predict CASP. However, many of these radiographic characteristics were dynamic, and we did not capture their changes over time in our analyses. Prospective analyses, powered to capture differences in RASP and CASP should be considered.

CONCLUSION

Our analysis of a highly controlled, robust data set supported the increased incidence or risk of RASP when patients were treated with ACDF rather than CDA. The mechanism of ACDF that causes patients to experience more RASP than those treated with CDA is not understood, but PROs remain unaffected by these differences. Identifying factors that increase the risk of RASP also remain debated and will be difficult to resolve with the inconsistencies in identification methods and rates of RASP.

REFERENCES


32. Nunley PD, Jawahar A, Cavanaugh DA, Gordon CR, Kerr EJ, Utter PA. Symptomatic adjacent segment disease after...


**Disclosures and COI:** Pierce Nunley receives royalties from K2M, Zimmer Biomet, Camber Spine, and Integrity Spine; is on the speakers bureau for K2M, Zimmer Biomet, Spineology, and Camber Spine; and is a paid consultant for K2M, Zimmer Biomet, Spineology, Vertiflex, Camber Spine, Integrity Spine, and Centinel Spine. He has stock or stock options in Paradigm, Spineology, and Camber Spine. Peter Campbell is a paid consultant for 4Web and Titan Spine. The other authors received no funding for this study and report no conflicts of interest.

**Corresponding Author:** Pierce D. Nunley, 1500 Line Ave, Suite 200, Shreveport, LA 71101. Phone: (318) 629-5555; Fax: (318) 629-5556; Email: pnunley@louisianaspine.org.

Published 30 June 2020
This manuscript is generously published free of charge by ISASS, the International Society for the Advancement of Spine Surgery. Copyright © 2020 ISASS. To see more or order reprints or permissions, see http://ijssurgery.com.