Randomized Trial

Comparison of Dysphagia Between Cervical Artificial Disc Replacement and Fusion

Data From a Randomized Controlled Study With Two Years of Follow-up

Martin Skeppholm, MD,*† and Claes Olerud, MD, PhD‡

Study Design. Prospective randomized controlled trial.

Objective. To determine and explain any differences in self-reported dysphagia between patients treated with artificial disc replacement and anterior cervical decompression and fusion (ACDF).

Summary of Background Data. Dysphagia after anterior cervical spine surgery has in previous studies been evaluated regarding different influencing factors. Surgical technique, number of treated levels, and type of implant has been shown to be of possible importance.

Methods. One hundred thirty-six patients from a randomized controlled trial between artificial disc replacement and ACDF in 1 or 2 surgical levels were evaluated regarding dysphagia. Evaluation was done with the dysphagia short questionnaire preoperatively, at 4 weeks, 3 months, and 1 and 2 years postoperatively. Reconstruction in the artificial disc replacement group was performed with the Discover artificial disc. Bone graft and anterior plating was used in the ACDF group. Type of implant was blinded to the patients and the surgeon until time of implantation.

Results. Demographics and dysphagia short questionnaire levels were similar in both groups preoperatively. At 4 weeks of follow-up postoperatively, dysphagia was significantly higher in both groups than baseline levels, $P < 0.01$. No significant differences were seen between the groups until follow-up at 2 years, which showed significantly higher dysphagia short questionnaire levels in the ACDF group, $P = 0.04$. The difference was statistically significant in both patients treated with 1- and 2-level surgery, $P = 0.029$ and $P = 0.032$, respectively. A logistic regression model showed a stronger association to type of implant than to number of surgical levels. Duration of surgery was highly associated to number of surgical levels but did not differ significantly between types of implant.

Conclusion. Long-term postoperative dysphagia could be explained by bulk of implant or decreased motion in the cervical spine. However, it is doubtful if differences between the groups in this study can be interpreted as a clinically important difference.

Key words: dysphagia, swallowing, cervical, artificial disc replacement, complications, randomized, spine surgery, anterior plating, bone graft, bulk, motion, nerve injury.

Level of Evidence: 2

Spine 2013;38:E1507–E1510

Dysphagia is probably one of the most common unwanted side effects of anterior cervical spine surgery, at least in the short term. Several prospective studies have shown that the problems in a majority of cases are transient, and also that dysphagia in some cases is present even before surgery.1–5 There are also reports on patients with more severe and protracted problems.6–8 The cause of long-term postoperative dysphagia is not fully understood, but conceivable factors such as bulk of implants and decreased cervical motion has been proposed in previous studies.9,10 Both factors could contribute to a difference in dysphagia between anterior cervical decompression and fusion (ACDF) and artificial disc replacement (ADR). In this part of a randomized study between ACDF and ADR, the aims were to determine if there was a difference in self-reported dysphagia between the groups and also if the number of surgical levels affected the results.

PATIENTS AND METHODS

One hundred thirty-six patients, 73 females and 63 males, with a mean age of 46.7 (standard deviation [SD], 6.7) years were enrolled in a multicenter study performed at 3 spine clinics in Sweden. The inclusion criteria for the randomized controlled trial was at least 3 months of radiicular arm pain with correlating magnetic resonance imaging findings at 1 or 2 cervical levels in patients aged between 25 and 60 years. The exclusion criteria were previous surgery in the cervical spine, severe
myelopathy, other cervical spine pathology, known hypersensitivity to implants or drugs used in the study, drug abuse, or dementia. The evaluation of dysphagia was performed with the dysphagia short questionnaire (DSQ), which has been validated in a previous study.\(^3\) All patients were operated with an anterior standard approach at 1 or 2 levels between C3–Th1, with a distribution as shown in Table 1. The randomization process was performed with closed and sealed envelopes. The content was not shown for the surgeon until it was time for reconstruction of the anterior column, to minimize the risk for adapting the surgical technique to the different implants. Intubation cuff pressure was measured and adjusted during surgery, and self-retaining retractors were not removed until time for wound closure. The Discover artificial disc (DePuy Spine, Raynham, MA) was used in the ADR group, whereas reconstruction in the ACDF group was performed with iliac crest bone grafting and fixation with an anterior plate of the surgeon’s preference. All patients received both oral and written information about the study before consent and inclusion. The study was approved by the regional ethics committee in Stockholm (2006/1266–31/3).

STATISTICS

Sample size was initially calculated on the basis of the primary outcome variable in the study, which is Neck Disability Index. With a given sample size of 16 patients in the smallest group (2-level ACDF) and \(\alpha = 0.05\), power was calculated post hoc to be 90%. The data used in the power analysis were taken from a previous study in which standard deviation and intraclass correlation was determined.\(^3\) The intraclass correlation was used to calculate standard error of measurement (SEM) and is defined by \(SEM = \sqrt{(1 - \text{intraclass correlation})}\).\(^11\) The smallest real difference (SRD) we wanted to detect on a group level is defined by \(SRD = 1.96 \times SEM \times \sqrt{2}\).\(^12\) SRD was used as an estimate for the minimal clinical important difference. The 2 groups were compared with parametric and nonparametric descriptive statistics. The Wilcoxon test for matched pairs was used for comparison within the groups. Kruskal-Wallis analysis of variance for repeated measurements was used for calculations of differences between the groups at different times of follow-up. The Fisher exact test was used for comparison of the groups divided into level of surgery. The variables group (ADR vs. ACDF) and level (1 vs. 2) were analyzed in a logistic regression model. For this analysis we dichotomized the DSQ such that values less than 4 was labeled as “less severe” and higher values as “more severe.” A DSQ score of 4 or more was chosen as a cut off because of our previous experience in clinical severity of dysphagia. The limit has also been used in a previous study of dysphagia using DSQ as an evaluation tool. \(T\) test was used to compare the duration of surgery (DOS) in the different groups. \(P\) values less than 0.05 were considered as statistically significant. Statistica software version 11 (StatSoft Inc., Tulsa, OK) was used for all calculations.

RESULTS

Both treatment groups were comparable in DSQ level and in demographic variables preoperatively, with data shown in Table 2. SEM and SRD were calculated to be 0.76 and 2.1, respectively. Compared with the preoperative levels, both the ADR and the ACDF group reported significantly higher levels of dysphagia 4 weeks postoperatively, \(P < 0.01\). After the first follow-up at 4 weeks, the DSQ levels declined over time in both groups as shown in Figure 1. There was no statistical difference in DSQ between the ADR and the ACDF groups during the first year of follow-up. At the 2-year follow-up, there was a statistical significant difference with the ADR group having a lower DSQ value, \(P = 0.04\) (Table 3) than the ACDF group. The difference in DSQ value between the groups was valid for both 1- and 2-level procedures, \(P = 0.029\) and \(P = 0.032\), respectively. At 4 weeks of follow-up, patients in the ADR group who got 1-level procedures showed lower DSQ value than those who got 2-level procedures, \(P = 0.03\). The difference in DSQ between the 1- and 2-level procedures could not be seen at any other time of follow-up in the ADR group, and not at any time in the ACDF group. The logistic regression analysis of the results at the 2-year follow-up showed an odds ratio of 0.019 for ADR compared with ACDF regarding more severe dysphagia, and this was statistically significant, \(P = 0.02\). One-level procedures compared with 2-level

**TABLE 1.** Distribution of Surgical Levels Between the Groups

<table>
<thead>
<tr>
<th>Levels</th>
<th>ADR, n (%)</th>
<th>ACDF, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C3–C4</td>
<td>0 (0)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>C4–C5</td>
<td>3 (4)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>C5–C6</td>
<td>26 (34)</td>
<td>21 (35)</td>
</tr>
<tr>
<td>C6–C7</td>
<td>24 (32)</td>
<td>21 (35)</td>
</tr>
<tr>
<td>C7–Th1</td>
<td>1 (1)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>C4–C6</td>
<td>3 (4)</td>
<td>2 (3)</td>
</tr>
<tr>
<td>C5–C7</td>
<td>19 (25)</td>
<td>14 (23)</td>
</tr>
</tbody>
</table>

ADR indicates artificial disc replacement; ACDF, anterior cervical decompression and fusion.

**TABLE 2.** Presentation of the Demographics in Both Groups

<table>
<thead>
<tr>
<th></th>
<th>ADR, n = 76</th>
<th>ACDF, n = 60</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSQ (range)</td>
<td>0 (0–8)</td>
<td>0 (0–8)</td>
<td>0.91</td>
</tr>
<tr>
<td>VAS neck (range)</td>
<td>62 (0–100)</td>
<td>60 (3–90)</td>
<td>0.61</td>
</tr>
<tr>
<td>Smokers, n (%)</td>
<td>23 (31)</td>
<td>19 (32)</td>
<td>0.85</td>
</tr>
<tr>
<td>2 levels, n (%)</td>
<td>22 (29)</td>
<td>16 (27)</td>
<td>0.63</td>
</tr>
<tr>
<td>Age (SD)</td>
<td>46.5 (6.7)</td>
<td>46.9 (6.8)</td>
<td>0.72</td>
</tr>
<tr>
<td>Females, n (%)</td>
<td>41 (54)</td>
<td>32 (52)</td>
<td>0.86</td>
</tr>
</tbody>
</table>

DSQ indicates dysphagia short questionnaire; ADR, artificial disc replacement; ACDF, anterior cervical decompression and fusion; VAS, visual analogue scale; SD, standard deviation.
procedures generated an odds ratio of 0.2 but not statistically significant, $P = 0.14$ (Table 4). There were no differences in mean DOS between the ADR group and the ACDF group, 125 minutes (SD, 42.6) compared with 141 minutes (SD, 38.0), $P = 0.18$. Average DOS in 1-level procedures was 122 minutes (SD, 35.9) compared with 160 (SD, 41.7) in the 2-level procedures, $P < 0.01$.

**DISCUSSION**

Several different factors may influence postoperative dysphagia after anterior cervical spine surgery. Various proposals have been made in previous studies with the aim to explain how swallowing is affected by surgery in this area. Possible causes for short-term dysphagia could be postoperative swelling of soft tissues and the esophageal wall, although a correlation in previous studies has not been clearly shown. Other studies have shown less swelling and dysphagia when steroids were administrated during surgery, indicating that an inflammatory reaction could contribute to the symptoms. Risk factors for more prolonged dysphagia has also been proposed in previous studies; multilevel surgery, long DOS, sex (females), bulk of implants, malalignment, and nerve damage. The question of whether patients undergoing surgery with ADR have less dysphagia compared with patients undergoing ACDF has previously been studied. McAfee et al showed less dysphagia in the ADR group both after short time follow-up as well as after 1 and 2 years. Two hundred fifty-one patients were included and were evaluated with the Bazaz score. One possible explanation for the different outcomes, according to the authors, could be a less traumatic surgical technique in the patients treated with ADR. Surgical approach and also whether specific surgical levels affects the outcome have previously been studied. Some doubts can be raised to the results of the study by McAfee et al because the evaluation of dysphagia was performed with a nonvalidated instrument, and to the fact that 16% of the patients had previous surgery. In contrast to the study by McAfee et al, our study design with randomization after decompression but before reconstruction, minimizes the risk for differences in pressure on esophagus and nerve damages which both theoretically can have an impact on dysphagia. Theoretically, there may be a difference in pressure and damage to soft tissue between 1- and 2-level surgery. However, significant higher levels of dysphagia could only be established in the ADR group at the first occasion of follow-up but not thereafter. We could not see any statistical differences between the treatment groups until 2 years after surgery, although the absolute values in the ACDF group were higher at all times of follow-up. Because both groups were treated similarly during surgery, it is more likely that dysphagia in a longer term is associated to the effects of the implants or the number of surgical levels. The logistic regression analysis showed a stronger association to the type of implant than to the number of surgical levels, and the significant difference in DSQ at 2 years was found in both the 1-and 2-level group. This finding indicates that the risk for higher levels of dysphagia in the long term is reduced in the ADR group. An analysis of DOS as risk factor was not performed because this factor was strongly associated to the number of surgical levels. Therefore, 2 factors affecting long-term dysphagia can be highlighted in this study. First, decreased motion of the cervical spine, and second, bulk of anterior plate fixation. One limitation of this study is that we cannot determine which 1 of these 2 factors contributes more to the development of long-term dysphagia in the ACDF group. It is also conceivable that altered alignment may contribute to disturbances in the swallowing mechanism in the ACDF group. However, these factors were not analyzed in this study. Although we found a statistically significant difference between the groups at the 2-year follow-up, the

---

TABLE 4. Odds Ratio for Higher Levels of DSQ (4 or Higher) Tested With Logistic Regression

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>OR</th>
<th>95% CI</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels (1 vs. 2 levels)</td>
<td>0.2</td>
<td>0.02–1.7</td>
<td>0.14</td>
</tr>
<tr>
<td>Group (ADR vs. ACDF)</td>
<td>0.019</td>
<td>0.001–0.53</td>
<td>0.02</td>
</tr>
</tbody>
</table>

**Adapted from** Copyright © 2013 Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
differences in DSQ on a group level were unobtrusive. The calculated SRD cannot directly be regarded as the minimal clinical important difference for the DSQ but can provide an indication of the approximate value. Taking this into account, the statistically significant difference on group level cannot uncritically be transferred to or interpreted as a clinically significant difference.

CONCLUSION
The findings in this study indicate that prolonged postoperative dysphagia could be explained by factors such as bulk of implants and decreased motion of the cervical spine. The results should be interpreted with caution because of uncertainty in clinically significant differences between the investigated groups. The number of surgical levels and DOS could not be associated with the outcome in the long term.

Key Points
- Postoperative dysphagia after cervical spine surgery is probably explained by several different factors.
- Postoperative dysphagia in the short and long term may have different causes.
- Swallowing performance is dependent of motion and alignment in the cervical spine.
- Bulk of implant may affect swallowing performance negatively in the long term.

Acknowledgment
The authors thank Eva Gulle for excellent data handling.

References