Heterotopic ossification following single-level anterior cervical discectomy and fusion: results from the prospective, multicenter, historically controlled trial comparing allograft to an optimized dose of rhBMP-2

Paul M. Arnold, MD,1 Karen K. Anderson, BS,1 Abdulhafez Selim, MD, PhD, MBA,2 Randall F. Dryer, MD,3 and J. Kenneth Burkus, MD4,5

1Department of Neurosurgery, University of Kansas Medical Center, Kansas City, Kansas; 2Medtronic Spinal and Biologics, Memphis, Tennessee; 3Central Texas Spine Institute, Austin, Texas; 4The Hughston Clinic; and 5Wilderness Spine Services, Columbus, Georgia

OBJECTIVE Heterotopic ossification (HO) has been reported following total hip, knee, cervical, and lumbar arthroplasty, as well as following posterolateral lumbar fusion using recombinant human bone morphogenetic protein–2 (rhBMP-2). Data regarding HO following anterior cervical discectomy and fusion (ACDF) with rhBMP-2 are sparse. A subanalysis was done of the prospective, multicenter, investigational device exemption trial that compared rhBMP-2 on an absorbable collagen sponge (ACS) versus allograft in ACDF for patients with symptomatic single-level cervical degenerative disc disease.

METHODS To assess differences in types of HO observed in the treatment groups and effects of HO on functional and efficacy outcomes, clinical outcomes from previous disc replacement studies were compared between patients who received rhBMP-2/ACS versus allograft. Rate, location, grade, and size of ossifications were assessed preoperatively and at 24 months, and correlated with clinical outcomes.

RESULTS Heterotopic ossification was primarily anterior in both groups. Preoperatively in both groups, and including osteophytes in the target regions, HO rates were high at 40.9% and 36.9% for the rhBMP-2/ACS and allograft groups, respectively (p = 0.350). At 24 months, the rate of HO in the rhBMP-2/ACS group was higher than in the allograft group (78.6% vs 59.2%, respectively; p < 0.001). At 24 months, the rate of superior-anterior adjacent-level Park Grade 3 HO was 4.2% in both groups, whereas the rate of Park Grade 2 HO was 19.0% in the rhBMP-2/ACS group compared with 9.8% in the allograft group. At 24 months, the rate of inferior-anterior adjacent-level Park Grade 2/3 HO was 11.9% in the rhBMP-2/ACS group compared with 5.9% in the allograft group. At 24 months, HO rates at the target implant level were similar (p = 0.963). At 24 months, the mean length and anteroposterior diameter of HO were significantly greater in the rhBMP-2/ACS group compared with the allograft group (p = 0.033 and 0.012, respectively). Regarding clinical correlation, at 24 months in both groups, Park Grade 3 HO at superior adjacent-level disc spaces significantly reduced range of motion, more so in the rhBMP-2/ACS group. At 24 months, HO negatively affected Neck Disability Index scores (excluding neck/arm pain scores), neurological status, and overall success in patients in the rhBMP-2/ACS group, but not in patients in the allograft group.

CONCLUSIONS Implantation of rhBMP-2/ACS at 1.5 mg/ml with polyetheretherketone spacer and titanium plate is effective in inducing fusion and improving pain and function in patients undergoing ACDF for symptomatic single-level cervical degenerative disc disease. At 24 months, the rate and dimensions (length and anteroposterior diameter) of HO were higher in the rhBMP-2/ACS group. At 24 months, range of motion was reduced, with Park Grade 3 HO in both treatment groups. The impact of Park Grades 2 and 3 HO on Neck Disability Index success, neurological status, and overall

ABBREVIATIONS ACDF = anterior cervical discectomy and fusion; ACS = absorbable collagen sponge; AE = adverse event; ALD = adjacent-level disease; ALOD = adjacent-level ossification development; AP = anteroposterior; DDD = degenerative disc disease; HO = heterotopic ossification; IDE = Investigational Device Exemption; NDI = Neck Disability Index; PEEK = polyetheretherketone; rhBMP-2 = recombinant bone morphogenetic protein–2; ROM = range of motion.


INCLUDE WHEN CITING Published online April 28, 2016; DOI: 10.3171/2016.1.SPINE15798.
Heterotopic ossification (HO) is a common complication following total hip and knee arthroplasty.\textsuperscript{13,17,22,35} Heterotopic ossification has also been reported in total disc arthroplasties in cervical\textsuperscript{5,8–12,15,16,18,21,24–26,29,32,37,39,41} and lumbar regions\textsuperscript{20,33,38} and in posterolateral lumbar spine fusion\textsuperscript{1,6} following the use of recombinant human bone morphogenetic protein–2 (rhBMP-2).

Recombinant human bone morphogenetic protein–2 is a recombinant osteoinductive protein that has been shown to induce bone formation between vertebrae when used as an implant with a suitable carrier under specific circumstances and at a specific concentration.\textsuperscript{3,27,28} Data regarding HO after anterior cervical disectomy and fusion (ACDF) with rhBMP-2 are sparse. In this study, we characterized HO in rhBMP-2 and allograft groups by rate, location, grade, size, and functional effects on range of motion (ROM), safety, and efficacy.

Cost comparison and evaluation of bone overgrowth or other early or late complications following the use of rhBMP-2 in ACDF were beyond the scope of this study. The study objectives were focused on the status of HO and clinical implications of HO and could pave the way for improved device designs. Clinical trial registration no.: IDE# G060021; data compared with pooled data from control arms of IDE# G010188/ NCT00642876 and IDE# G000123/NCT00437190 (www.clinicaltrials.gov).

http://thejns.org/doi/abs/10.3171/2016.1.SPINE15798

KEY WORDS absorbable collagen sponge; anterior cervical disectomy and fusion; degenerative disc disease; heterotopic ossification; INFUSE bone graft; recombinant human bone morphogenetic protein–2

Methods

Study Design

This study was a subanalysis of the main prospective, multicenter, nonrandomized, historically-controlled FDA Investigational Device Exemption (IDE) trial comparing the safety and effectiveness of rhBMP-2 delivered on an absorbable collagen sponge (ACS) (INFUSE Bone Graft, Medtronic Spinal and Biologics) with a polyetheretherketone (PEEK) interbody spacer and titanium anterior cervical plate in facilitating spinal fusion in patients with symptomatic single-level cervical degenerative disc disease (DDD) from C-3 to C-7 (Study\# P05–04, IDE\# G060021, Pivotal/Phase III).

This subanalysis evaluated differences between specific grades and/or locations of HO in the rhBMP-2/ACS and allograft groups, and evaluated if specific grades and/or locations of HO had any effect on safety and efficacy outcomes. Safety and effectiveness data for the rhBMP-2/ACS group were compared with historical pooled data from the control arms (patients who underwent single-level ACDF) of 2 studies evaluating single-level cervical disc arthroplasty versus ACDF in the treatment of symptomatic cervical DDD. These studies were: 1) the Pivotal IDE Study of the PRESTIGE Cervical Disc Device (IDE\# G010188, NCT00642876), and 2) the Pivotal IDE Study of the BRYAN Cervical Disc Prosthesis (IDE\# G000123, NCT00437190). The relevant institutional review boards or ethics committees approved the study protocol, and patients’ informed consent was obtained.

Control patients underwent single-level ACDF with cortical ring allograft interbody spacers and local bone reaming and the ATLANTIS Cervical Plate System (Medtronic Sofamor Danek USA, Inc.). The posterior longitudinal ligament was removed in all cases. Perioperative steroids, physical activity, and use of external orthosis were at the discretion of the operating surgeon.

The INFUSE Bone Graft/ACDF pivotal IDE trial involved all patients receiving 1.5 mg/ml rhBMP-2 in a PEEK interbody spacer implanted via an open ACDF. The total dose of INFUSE implanted was 0.6 mg (0.4 ml on an 0.4-cm\textsuperscript{3} ACS) or 1.05 mg (0.7 ml on an 0.7-cm\textsuperscript{3} ACS), depending on the size of the PEEK spacer chosen by the operating surgeon.

Data were collected preoperatively, intraoperatively, at discharge, 6 weeks postoperatively, and 3, 6, 12, and 24 months postoperatively. Patients in the rhBMP-2/ACS group were followed until the last patient completed the 24-month visit, whereas patients in the allograft group were followed beyond 24 months. Additional details are available at http://www.clinicaltrials.gov.

Heterotopic ossification in the target-level region was evaluated preoperatively and at 24 months postoperatively with neutral anteroposterior (AP) and lateral radiographs and dynamic flexion-extension radiographs. Sagittal plane angulation was measured on dynamic lateral radiographs. Kissing or bridging osteophytes were documented as a single HO. If the HO spanned several levels, the HO was recorded as 1. If an HO extended from outside the target level to adjacent levels above or below, the location was recorded from the superior-most or inferior-most level outside and within the region. Heterotopic ossification in the target-level region was described as osteophytic, well or poorly margined, and any degenerative changes at the superior and inferior adjacent endplates were noted.

Severity of HO was graded using the Park classification system: Grade 0 (no HO); Grade 1 (mild; ossification across < 50% of adjacent disc space); Grade 2 (moderate; ossification across ≥ 50% of adjacent disc space); and Grade 3 (severe; HO completely bridging adjacent disc space)\textsuperscript{31} (Fig. 1). Ossifications of Grades 0–2 are better
described as osteophytes, whereas Grades 3 and 4 could be classified as HO. Because we were using the Park classification for HO, the use of HO terminology was more appropriate. However, because HO was not graded with the Park system in patients receiving allograft in the historical studies, the FDA permitted Biomedical Systems to evaluate HO in both groups. For assessment centralization, radiographs at each site were read by 3 groups of independent reviewers. Radiographs were assigned randomly to each team of 2 primary readers and 1 adjudicator, for a total of 9 reviewers. Heterotopic ossification dimensions (length, AP diameter, and size) were calculated by the mean of the 2 largest readings among the 3 readers. If > 1 HO was noted, the largest HO (distance from superior to inferior × AP dimension) was used. Although this was an industry-sponsored study and a coauthor is an employee of the sponsor, the data were independently analyzed by the Vanderbilt University Medical Center Biostatistics Collaboration Center.

Three HO subgroups (Park Grades 0/1, 2, and 3) were derived according to the adjudicated Park classification at the superior adjacent disc, inferior adjacent level, and the target implant level. Patients were grouped based on the highest grade among the 3 locations and then compared within the same group.

Angular ROM and translation at the level above the target level were summarized within each group and compared by the 3 HO subgroups at the superior adjacent disc. Similarly, angular ROM and translation at the level below the target level were compared by the 3 HO subgroups at the inferior adjacent disc. The p values for comparing the 3 subgroups were presented for each group by using ANCOVA, with the preoperative value as the covariate. Adjacent-level angular and translatory motion data were pooled from Biomedical Systems for the rhBMP-2/ACS group and from SYNARC, Inc., for the allograft group.

The effectiveness outcomes of arm and neck pain scores, Neck Disability Index (NDI) scores, NDI success, neurological success, and overall success were summarized for the 3 HO subgroups. Anterior and posterior HOIs were compared separately. All p values were based on ANCOVA by using a corresponding preoperative score as the covariate except for neurological success and overall success, for which Fisher’s exact test was used. Overall success was defined the same way in both groups: fusion, NDI, and neurological success; no implant-associated or implant/procedure-associated serious adverse events (AEs); and no additional surgeries or interventions.

Cumulative rates of AEs and secondary surgeries at the index level at pre- and post-6-months were summarized by the 3 HO subgroups. The pre–6-month rate included onset during or before the 3-month AE window, whereas the post–6-month rate included onset during the 6-, 12-, and 24-month AE windows.

A fixed value of 0.10, the threshold commonly used by the FDA to determine noninferiority, was used as the noninferiority margin for assessing all noninferiority hypotheses regarding safety and operative measurements. For AEs and additional surgeries or interventions, a superiority hypothesis was used. The primary data set consisted of all patients who received rhBMP-2/ACS and underwent ACDF. Primary statistical comparisons were based on the observed data, and missing data due to lost follow-ups were not imputed.

There was no predefined analysis for HO because it was not a study end point. Nonetheless, for the presence of any preoperative HO, the p value was from Fisher’s exact test, and a p value for HO at 24 months was from logistic regression, with propensity score and any preoperative HO as the covariates. For the assessment of preoperative HO by grade at different locations, the p value was from Fisher’s exact test; for HO at 24 months, the p value was from logistic regression, comparing Park Grade 2/3 rates between the groups by using a propensity score and preoperative HO at the superior adjacent disc (Grades 0/1 vs 2/3) as the covariates. For the relationship between different locations of HO and overall success in patients in the allograft group, Fisher’s exact test was used. For the mean ROM by HO grade at different locations, p values

TABLE 1. Summary of any ossification preoperatively

<table>
<thead>
<tr>
<th>Value*</th>
<th>rhBMP-2/ACS, n = 224 (%)</th>
<th>Allograft, n = 486 (%)</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any preop ossification</td>
<td>90/220 (40.9)</td>
<td>163/442 (36.9)</td>
<td>0.350</td>
</tr>
<tr>
<td>No. of preop ossifications</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>130 (59.1)</td>
<td>279 (63.1)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>37 (16.8)</td>
<td>80 (18.1)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>30 (13.6)</td>
<td>53 (12.0)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>16 (7.3)</td>
<td>23 (5.2)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5 (2.3)</td>
<td>6 (1.4)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1 (0.5)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1 (0.5)</td>
<td>1 (0.2)</td>
<td></td>
</tr>
</tbody>
</table>

* Established using Fisher’s exact test.
Results

The original prospective, historically controlled trial began in June 2007. The last surgery was performed in January 2010, and the last 24-month follow-up visit was in December 2011. There were 224 patients in the rhBMP-2/ACS group and 486 patients in the allograft group, for a total of 710 patients. At 24 months, the follow-up rate for patients in the rhBMP-2/ACS group was 83.93% (188 of 224) compared with 86.83% (422 of 486) for patients in the allograft group. The surgeries were performed by 18 principal investigators at 17 sites in the US. No single site contributed more than 20% of all patients. There were no statistically significant differences in age, sex, smoking, litigation, or work status between the 2 groups. The fusion success rate for the rhBMP-2/ACS group at 24 months was shown to be significantly higher than that in the allograft group (99.4% compared with 87.2%; p = 0.002).

Pseudarthrosis was the most common device-related AE, and it was higher in the allograft group. Whereas 1 pseudarthrosis event was reported in the rhBMP-2/ACS group, 40 events were reported in 40 subjects in the allograft group.

Heterotopic ossification was documented in 7 vertical locations from superior to inferior in relation to the implant, and in 4 locations from anterior to posterior. No HO was observed at lateral locations in either group at any time; all HO on radiographs was either anterior or posterior. Preoperatively (including osteophytes in the target regions), there was no significant difference in the rate of HO

<table>
<thead>
<tr>
<th>Variable</th>
<th>rhBMP-2/ACS, n = 224 (%)</th>
<th>Allograft, n = 486 (%)</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any ossification at 24 mos</td>
<td>132/168 (78.6)</td>
<td>241/407 (59.2)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

No. of ossifications at 24 mos

<table>
<thead>
<tr>
<th>No.</th>
<th>rhBMP-2/ACS (%)</th>
<th>Allograft (%)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>36 (21.4)</td>
<td>166 (40.8)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>54 (28.1)</td>
<td>144 (35.4)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>51 (30.4)</td>
<td>69 (17.0)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>22 (13.1)</td>
<td>24 (5.9)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4 (2.4)</td>
<td>4 (1.0)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1 (0.6)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0 (0.5)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
</tbody>
</table>

* Logistic regression using propensity score and any ossification (yes/no) preoperatively as the covariates.
between groups (p = 0.350), although the rates were high, at 40.9% and 36.9% for the rhBMP-2/ACS and allograft groups, respectively (Table 1). At 24 months, the rate of HO was higher than the preoperative rate in both groups; however, the rate of HO in the rhBMP-2/ACS group was significantly higher (78.6% vs 59.2%) than in the allograft group after adjusting for preoperative HO and propensity score (p < 0.001) (Table 2).

Preoperatively, the rate of HO at the superior adjacent disc was similar in both groups (p = 0.246; Fisher’s exact test comparing Grades 0/1 vs 2/3). At 24 months, the rate of superior-anterior adjacent-level HO was higher in both groups compared with preoperative rates, but the difference between the groups was not statistically significant after adjusting for propensity score and preoperative HO status (p = 0.322; comparing Grades 0/1 vs 2/3). Nonetheless, at 24 months, the rate of superior-anterior Park Grade 2 HO was 19.0% among patients in the rhBMP-2/ACS group compared with 9.8% among patients in the allograft group; the rate of superior-anterior adjacent-level Park Grade 3 HO was 4.2% in both groups (Table 3).

Preoperatively, the rate of HO at the inferior adjacent disc was similar in both groups (p = 0.117; Fisher’s exact test comparing Grades 0/1 vs 2/3). At 24 months, the rate of inferior-anterior adjacent-level HO was higher in both groups compared with preoperative rates, but the difference between the groups was not statistically significant after adjusting for propensity score and preoperative HO status (p = 0.360; comparing Grades 0/1 vs 2/3). Nonetheless, at 24 months, the rate of inferior-anterior adjacent-level Park Grade 2/3 HO was 11.9% among patients in the rhBMP-2/ACS group compared with 5.9% among patients in the allograft group (Table 3).

At 24 months, in both groups, Park Grade 3 HO at superior adjacent-level disc spaces significantly decreased angular ROM, more so in the rhBMP-2/ACS group (Fig. 2). Preoperatively, the rate of HO at the target implant level was similar in both groups (p = 0.114). At 24 months, HO rates at this level were still similar (p = 0.963). Only 1 patient in the rhBMP-2/ACS group developed HO (Grade 2). No patients in the allograft group developed HO (Table 4).

Preoperatively, there was no significant difference in average dimensions of HO (length, AP diameter, and size) between the groups. At 24 months, the mean length and AP diameter of HO were significantly greater in the rhBMP-2/ACS group compared with the allograft group, after adjusting for propensity score and corresponding preoperative dimensions (p = 0.033 and 0.012, respectively) (Table 5; Figs. 3 and 4).

At 24 months, HO negatively affected NDI success (excluding neck or arm pain scores), neurological status, and overall success among patients in the rhBMP-2/ACS group, but a similar negative relationship was not observed among patients in the allograft group. It is not certain whether that relationship was real or merely a chance finding, because statistical analyses of small subgroups are vulnerable to this phenomenon (Tables 6 and 7; Figs. 5–10).

### Table 4. Ossifications at target implant level by Park grade

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Park Grade</th>
<th>rhBMP-2/ACS, n = 224 (%)</th>
<th>Allograft, n = 486 (%)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preop</td>
<td>0</td>
<td>202 (91.8)</td>
<td>422 (95.5)</td>
<td>0.114†</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>6 (2.7)</td>
<td>9 (2.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>12 (5.5)</td>
<td>11 (2.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>24 mos postop</td>
<td>0</td>
<td>167 (99.4)</td>
<td>407 (100.0)</td>
<td>0.963†</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1 (0.6)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
</tbody>
</table>

* Fisher’s test comparing Grade 0/1 versus Grade 2/3.
† Determined by comparing Grade 0/1 versus Grade 2/3, with logistics regression adjusting for preoperative HO and propensity score.

### Table 5. Summary of ossification size

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Variable</th>
<th>Mean (SD), Median</th>
<th>p Value*</th>
<th>p Value†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>rhBMP-2/ACS, n = 224</td>
<td>Allograft, n = 486</td>
<td></td>
</tr>
<tr>
<td>Preop</td>
<td>Length, mm</td>
<td>3.20 (4.16), 1.45</td>
<td>2.66 (3.50), 1.30</td>
<td>0.239</td>
</tr>
<tr>
<td></td>
<td>AP diameter, mm</td>
<td>1.87 (2.17), 0.90</td>
<td>1.74 (2.08), 0.95</td>
<td>0.429</td>
</tr>
<tr>
<td></td>
<td>Area size, mm²</td>
<td>14.96 (30.28), 2.50</td>
<td>12.16 (25.59), 2.36</td>
<td>0.369</td>
</tr>
<tr>
<td>24 mos postop</td>
<td>Length, mm</td>
<td>7.33 (5.95), 6.12</td>
<td>5.13 (5.76), 4.14</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>AP diameter, mm</td>
<td>3.84 (2.43), 3.92</td>
<td>2.88 (2.74), 2.92</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Area size, mm²</td>
<td>39.89 (47.60), 24.56</td>
<td>29.39 (58.29), 11.94</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

* Determined using unadjusted, Wilcoxon rank-sum nonparametric test.
† Determined using ANCOVA adjusting for propensity score and corresponding preoperative dimension (length, AP diameter, and size) score.

Discussion

The surgical implantation of INFUSE Bone Graft
(rhBMP-2/ACS) at a concentration of 1.5 mg/ml with a PEEK interbody spacer and titanium anterior plate is a safe and effective replacement for allograft bone in inducing fusion and improving pain and function in patients undergoing ACDF for symptomatic single-level cervical DDD from C-3 to C-7 (JK Burkus et al., unpublished data). Although higher superiority margins would be anticipated in multiple-level compared with single-level ACDF, evaluation of single-level ACDF had to be conducted first because of regulatory considerations, dose-finding, and other optimizations needed for multiple-level studies. In addition, the use of the INFUSE Bone Graft/PEEK interbody spacer instead of allograft bone avoids the potential risks of disease transmission and graft failure/fracture that are associated with allograft.

Several studies have examined various approaches to prevent or reduce HO formation, including medications, surgical techniques, and device design modifications.\(^2,7,14,19,30\) The use of perioperative NSAIDs has been shown to reduce the risk of developing HO by 59% over placebo.\(^14\) In our study, the use of NSAIDs was an exclusion criterion. Pradhan et al. evaluated 10 patients who underwent ACDF with fibular allograft and rhBMP-2/ACS. Fibrin glue was placed in the posterior aspect of the interbody space to seal it off from the spinal canal. At 3 months, there was no evidence of heterotopic bone in the canal.\(^34\) In our study, the rate of posterior HO was very low among patients who received allograft, and no posterior HO was detected in patients who received rhBMP-2/ACS even though fibrin glue was not used.

**FIG. 3.** Dimensions of HO before surgery. Values on the y axis are millimeters. A = area; D = diameter; INFUSE = rhBMP-2/ACS; L = length.

**FIG. 4.** Dimensions of HO at 24 months. INFUSE = rhBMP-2/ACS. Values on the y axis are millimeters. *From unadjusted, Wilcoxon rank-sum nonparametric test. **From ANCOVA adjusting propensity score and corresponding preoperative dimension (length, AP diameter, size) score.
Previous studies that addressed cervical HO excluded subjects with preoperative HO.\(^{15,40}\) The data in our study were adjusted to the preoperative HO rates, which enabled us to evaluate the effect of rhBMP-2/ACS on preexisting as well as new HO formation. At 24 months, both groups demonstrated a marked increase in HO rates compared with preoperative rates, suggesting that HO formation may have been triggered by the ACDF procedure, regardless of the device used.\(^{21}\) In fact, HO has developed after cervical procedures other than ACDF, such as disc replacement.\(^{4,18,29}\)

Park Grade 2/3 HO may severely impact or eliminate ROM, and decreased ROM has been proposed as a predisposing factor for adjacent-level ossification development (ALOD).\(^{21}\) In our study, at 24 months, Park Grade 3 HO at superior adjacent-level disc spaces significantly decreased angular ROM in both groups.

Uninstrumented fusion has shown notably lower rates of ALOD compared with instrumented fusion.\(^{15,31,40}\) Yang et al. found that only 5.5% of patients had ALOD after uninstrumented fusion.\(^{40}\) The authors recommend minimal stripping of the anterior longitudinal ligament and avoiding Caspar pins and anterior plates to reduce ALOD. In their review of 118 patients who underwent instrumented anterior fusion, Park et al. observed an ALOD rate of 59% in the cephalad adjacent disc spaces and 29% in the caudal adjacent disc spaces.\(^{31}\) Superior HO may be associated with 2-fold higher rates of ALOD when compared with inferior HO, even in uninstrumented fusion.\(^{18}\)

Another risk factor for ALOD is proximity of the plate to the adjacent disc space. Park et al. found an ALOD rate at the cephalad adjacent disc spaces of 67% when the distance was < 5 mm versus 24% at ≥ 5 mm, and a rate of 45% at < 5 mm versus 5% at ≥ 5 mm at the caudal adjacent disc spaces.\(^{31}\) Lee et al. used longer cranial and caudal screws

### TABLE 6. Effect of ossification on success in the allograft group at 24 months

<table>
<thead>
<tr>
<th>Variable</th>
<th>Grade 0/1</th>
<th>Grade 2</th>
<th>Grade 3</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDI</td>
<td>270/334</td>
<td>42/53</td>
<td>14/19</td>
<td>0.687</td>
</tr>
<tr>
<td>%</td>
<td>80.8</td>
<td>79.2</td>
<td>73.7</td>
<td></td>
</tr>
<tr>
<td>Neurological</td>
<td>290/334</td>
<td>42/54</td>
<td>15/19</td>
<td>0.15</td>
</tr>
<tr>
<td>%</td>
<td>86.8</td>
<td>77.8</td>
<td>78.9</td>
<td></td>
</tr>
<tr>
<td>Fusion</td>
<td>243/288</td>
<td>43/46</td>
<td>12/12</td>
<td>0.176</td>
</tr>
<tr>
<td>%</td>
<td>85.6</td>
<td>93.5</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>173/288</td>
<td>29/46</td>
<td>8/13</td>
<td>0.943</td>
</tr>
<tr>
<td>%</td>
<td>60.1</td>
<td>63.0</td>
<td>61.5</td>
<td></td>
</tr>
</tbody>
</table>

* Determined using Fisher’s exact test.
angled away to maximize the distance to the adjacent endplates, and found a significantly lower ALOD rate in the investigational group at both the cranial (42% vs 72%) and caudal (20% vs 42%) adjacent disc spaces.\textsuperscript{23} The plate proximity factor was not part of our study protocol, because the study was not designed to evaluate the impact of this potential confounding variable. Because all study patients underwent the same surgical procedure, the effect of endplate distance on HO was not anticipated to be a “fixed-effect” major confounding variable. Regarding adjacent-level disease (ALD), it is important to mention that in this study, only 3 cases of minimal ALD were observed in the rhBMP2/ACS group. One of the 3 cases had small calcifications at the inferior plate. In the allograft group, 12 cases with mild to moderate ALD were observed, and 2 of those cases were suggested to be related to HO.

One limitation of this study was the use of a historical (nonconcurrent) control group; thus, this study lacked the advantages of randomization. Because of the risk that the 2 treatment groups may not have been comparable, we sought to balance the covariates via propensity score methods. In addition, we sought to balance distribution of patient characteristics and limit or eliminate selection and temporal bias.

There may be a potential additive effect of the investigational device. This effect was noted in a meta-analysis of 5 noncervical studies, indicating a trend toward increased risk of HO development associated with rhBMP-2.\textsuperscript{36} We did not evaluate dose effect in our study because only 4 pa-

**FIG. 6.** Relationship between HO at 24 months and NDI success in the allograft group.

**FIG. 7.** Relationship between HO at 24 months and neurological success in the rhBMP-2/ACS group.
tients received the higher dose of 1.05 mg, compared with 220 patients who received 0.6 mg. It is possible that dose-finding studies may identify lower dose(s) of rhBMP-2 that will demonstrate lower HO rates without compromising effectiveness.

Conclusions

We characterized HO after ACDF in rhBMP-2/ACS and allograft groups by rate, location, grade, and size, and investigated the correlation of HO to overall success and clinical performance. Preoperative HO rates were similar in both groups. At 24 months, both groups had higher HO rates when compared with preoperative rates. Postoperative HO rates were higher in the rhBMP-2/ACS group. Preoperatively, there was no significant difference in average dimensions of HO between the groups. At 24 months, the mean length and AP diameter of HO were significantly greater in the rhBMP-2/ACS group compared with the allograft group. At 24 months, Park Grade 3 HO at superior adjacent-level disc spaces significantly decreased angular ROM in both treatment groups. At 24 months, HO grades had mixed impacts on NDI success (excluding neck or arm pain scores), neurological status, and overall success. All of these findings should be considered when assessing the risks and benefits of using rhBMP-2/ACS in ACDF.

Acknowledgments

Guorong Ma, PhD, Senior Clinical Research Director at Medtronic Spine and Biologics, performed statistical analyses. David F. Wootten, PhD, Director of Clinical Science at Medtronic...
ic Spine and Biologics, contributed to data interpretation and reviewed the manuscript for technical accuracy.

References


Disclosures
Sources of Support (funding, technical support, and corporate support): The sponsor of this FDA-approved IDE clinical trial was Medtronic Spinal and Biologics (Medtronic Sofamor Danek USA, Inc. [“Medtronic”], 1800 Pyramid Place, Memphis, TN 38132). The authors are consultants and/or clinical investigators for Medtronic Spinal and Biologics, the company which distributed the devices that were studied. Although this was an industry-sponsored study, and Dr. Selim is an employee of the sponsor, the data were independently analyzed by the Vanderbilt University Medical Center Biostatistics Collaboration Center. Potential Conflicts of Interest (and sources of funding): No author received any financial support for this work. Dr. Selim has a PhD in bone biology and diploma in applied biostatistics, and he contributed to the study design and data analysis. He is an employee of Medtronic. Dr. Arnold is a consultant for Medtronic, Stryker Spine, and FioMed. He has stock ownership in Z-Plasty, and has received sponsored or reimbursed travel from Stryker Spine and AOSpine North America. Dr. Burkus is a consultant for Zimmer Biomet and Medtronic. He owns patents, and receives royalties from Zimmer Biomet and Medtronic. He has received payment from Zimmer Biomet and Medtronic for development of educational presentations. He has received sponsored or reimbursed travel funds from Zimmer Biomet and Medtronic.

Author Contributions
Conception and design: all authors. Acquisition of data: all authors. Analysis and interpretation of data: all authors. Drafting the article: Anderson. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Arnold.

Correspondence
Paul M. Arnold, Department of Neurosurgery, Mail Stop 3021, University of Kansas Medical Center, 3901 Rainbow Blvd., Kansas City, KS 66160. email: parnold@kumc.edu.