Exploratory meta-analysis on dose-related efficacy and morbidity of bone morphogenetic protein in spinal arthrodesis surgery

*Christoph P. Hofstetter, MD, PhD,1 Anna S. Hofer, MD,2 and Allan D. Levi, MD, PhD2

1Department of Neurological Surgery, University of Washington, Seattle, Washington; and 2Department of Neurological Surgery and The Miami Project to Cure Paralysis, University of Miami Miller School of Medicine, Miami, Florida

OBJECT Bone morphogenetic protein (BMP) is frequently used for spinal arthrodesis procedures in an “off-label” fashion. Whereas complications related to BMP usage are well recognized, the role of dosage is less clear. The objective of this meta-analysis was to assess dose-dependent effectiveness (i.e., bone fusion) and morbidity of BMP used in common spinal arthrodesis procedures. A quantitative exploratory meta-analysis was conducted on studies reporting fusion and complication rates following anterior cervical discectomy and fusion (ACDF), posterior cervical fusion (PCF), anterior lumbar interbody fusion (ALIF), transforaminal lumbar interbody fusion (TLIF), posterior lumbar interbody fusion (PLIF), and posterolateral lumbar fusion (PLF) supplemented with BMP.

METHODS A literature search was performed to identify studies on BMP in spinal fusion procedures reporting fusion and/or complication rates. From the included studies, a database for each spinal fusion procedure, including patient demographic information, dose of BMP per level, and data regarding fusion rate and complication rates, was created. The incidence of fusion and complication rates was calculated and analyzed as a function of BMP dose. The methodological quality of all included studies was assessed according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. Data were analyzed using a random-effects model. Event rates are shown as percentages, with a 95% CI.

RESULTS Forty-eight articles met the inclusion criteria: ACDF (n = 7), PCF (n = 6), ALIF (n = 9), TLIF/PLIF (n = 17), and PLF (n = 9), resulting in a total of 5890 patients. In ACDF, the lowest BMP concentration analyzed (0.2–0.6 mg/level) resulted in a fusion rate similar to the highest dose (1.1–2.1 mg/level), while permitting complication rates comparable to ACDF performed without BMP. The addition of BMP to multilevel constructs significantly (p < 0.001) increased the fusion rate (98.4% [CI 95.4%–99.4%]) versus the control group fusion rate (85.8% [CI 77.4%–91.4%]). Studies on PCF were of poor quality and suggest that BMP doses of ≤2.1 mg/level resulted in similar fusion rates as higher doses. Use of BMP in ALIF increased fusion rates from 79.1% (CI 57.6%–91.3%) in the control cohort to 96.9% (CI 92.3%–98.8%) in the BMP-treated group (p < 0.01). The rate of complications showed a positive correlation with the BMP dose used. Use of BMP in TLIF had only a minimal impact on fusion rates (95.0% [CI 92.8%–96.5%] vs 93.0% [CI 78.1%–98.0%] in control patients). In PLF, use of ≥8.5 mg BMP per level led to a significant increase of fusion rate (95.2%; CI 90.1%–97.8%) compared with the control group (75.3%; CI 64.1%–84.0%, p < 0.001). BMP did not alter the rate of complications when used in PLF.

CONCLUSIONS The BMP doses used for various spinal arthrodesis procedures differed greatly between studies. This study provides BMP dosing recommendations for the most common spine procedures.

http://thejns.org/doi/abs/10.3171/2015.4.SPINE141086

KEY WORDS bone morphogenetic protein; dosage; fusion; meta-analysis; spinal; technique
More than 400,000 spinal fusion procedures are performed annually in the US. Autologous iliac crest bone graft is considered the gold standard for bone grafting material due to its superior osteoconductive, osteoinductive, and osteogenic properties. However, donor site morbidity as well as high demands of the quality of bone grafting material commanded by minimal access surgery and by long-segment instrumented fusions in deformity surgery, has fueled the quest for alternatives to existing bone graft substitutes. Bone morphogenetic protein (BMP), a prototypical osteoinductive protein belonging to the superfamily of transforming growth factor-beta, was first discovered in 1965 by Marshall R. Urist. Recombinant human bone morphogenetic protein–2 (rhBMP-2) has been tested clinically for use in spinal arthrodesis in several prospective, randomized clinical trials beginning in 1997. Clinical studies of lumbar and cervical fusion in humans have demonstrated consistent osteoinduction mediated by rhBMP-2–impregnated absorbable collagen sponges (INFUSE) placed into intervertebral spacers. In 2002, the US FDA approved the use of rhBMP-2 in combination with tapered, threaded, lordotic fusion cages (LT-CAGE; Medtronic Sofamor Danek) to enhance anterior lumbar spinal fusion after an FDA trial showed efficacy and safety. Recombinant human osteogenic protein–I was approved in 2004 to supplement revision posterior spinal fusions. Because early studies indicated that patients receiving rhBMP-2 achieved solid fusions without major adverse events in anterior lumbar interbody fusion (ALIF) procedures, rhBMP-2 has been used in an “off-label” fashion in more than 85% of spinal fusion cases.

The first concerns regarding the safety of rhBMP-2 emerged in cases of anterior cervical discectomy and fusion (ACDF). In particular, prevertebral swelling leading to dysphagia or in some cases life-threatening airway obstruction was reported in up to 27.5% of cases. In 2008, the FDA issued a statement warning spinal surgeons to avoid use of rhBMP-2 in the anterior cervical spine until further data were available regarding efficacy and safety. Since then, rhBMP-2 has been associated with multiple serious adverse effects including soft-tissue swelling, local inflammation, sterile cyst formation, osteolysis, implant migration, and ectopic bone formation as well as possible increased risk of malignancy in the high-dose AMPLIFY formulation (Medtronic Sofamor Danek). BMP dosing for different procedures varies considerably in the published literature, which indicates that there is great uncertainty regarding optimal dosing for specific spinal fusion procedures. The objective of the current meta-analysis was to analyze the effect of BMP dosing on successful arthrodesis and morbidity in the most commonly performed spinal fusion procedures.

Methods

Search Strategy

PubMed and The Cochrane Library databases were searched using combinations of the following key words: anterior cervical discectomy and fusion (ACDF), posterior cervical fusion (PCF), anterior lumbar interbody fusion (ALIF), transforaminal lumbar interbody fusion (TLIF), posterior lumbar interbody fusion (PLIF) and posterolateral lumbar fusion (PLF), and bone morphogenetic protein (BMP). The words were searched either in the title, abstract, or key words. A hand search of reference lists of obtained articles was also performed. The last search was conducted in August 2014. For data collection, analysis, and preparation of the manuscript, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement for reporting systematic reviews and meta-analyses was followed (Fig. 1; PRISMA checklist not published).

Inclusion and Exclusion Criteria

All studies that reported BMP-dependent fusion rates and/or perioperative morbidity were identified (n = 1356, Fig. 1). Studies were eligible if they reported BMP dose per fusion level and individual patient outcomes dependent on BMP dose. Control groups were only included if allograft or autograft was used. In studies that used ceramic bone extenders or alternative sources of BMP for control patients, only the data of the BMP treatment arm was extracted and included in our analysis. Only clinical studies with a minimum of 15 participants and 1 year of follow-up for fusion assessment were included. In vitro and animal studies, case reports, biomechanical or cadaveric studies, or articles in languages other than English or German were excluded. In cases of multiple articles from the same author(s) that clearly reported the same cohort of patients, we either included only the most recently published article, or we included both studies but separated the number of patients and outcome rates so that patient data were only included once in our analysis. Three authors were contacted to provide missing data necessary to fulfill inclusion criteria—only 1 author responded. Eventually, 48 articles met our inclusion criteria: ACDF (n = 7), PCF (n = 6), ALIF (n = 9), TLIF/PLIF (n = 17), and PLF (n = 9).

Data Extraction

Each article was analyzed and all pertinent information regarding patient and study characteristics, operative data, and results was abstracted by 2 investigators and confirmed by the senior author (A.L.). Reviewers were not blinded to authors, journal, or source of financial support. Any disagreements were discussed among the investigators until consensus was reached. Articles were grouped based on the type of spine procedure investigated. From the included studies, a database was created for each spinal fusion procedure, providing information about author(s); study design; year; type of procedure; number of levels fused; dose of BMP per level; number of patients in total, number of patients receiving BMP, and number of patients in the control group (if present); description of the surgical technique (instrumentation, type of interbody device like cage or structural allograft, placement of BMP, use of additional autograft or allograft); number of patients undergoing fusion assessment; and number and rate of patients showing fusion at last follow-up as
well as types, number, and rate of complications, and industry relationship. To allow for interstudy comparisons, complications were strictly defined for each procedure analyzed.

In ACDF, BMP-related prevertebral cervical soft-tissue swelling has been proposed to lead to postoperative dysphagia, dysphonia, and respiratory difficulties. However, these complications lack a standardized definition in the current literature and therefore an interstudy comparison was not feasible. Therefore, we included only the most severe complications that were reliably documented by the included studies (readmission to the hospital, reintubation and/or tracheostomy, or return to the operating room). For PCF, the included reports provide information regarding neurological (postoperative radiculopathies) and wound complications. Excessive neck pain was not considered to be a radicular complication. For ALIF, retrograde ejaculation, endplate resorption, and graft subsidence were reported. Retrograde ejaculation is typically assessed via questionnaires; only 1 study group performed a quantitative analysis of retrograde ejaculation. Studies on TLIF reported radiculopathy, ectopic/heterotopic bone formation, seroma, and wound infection. BMP-related radiculopathy corresponded to the dermatomal distribution of the TLIF window and occurred more than 10 days after the procedure in a patient who had previously been asymptomatic postoperatively. Cases of radiculopathy due to pseudarthrosis or malpositioned pedicle screws were not included in our analysis because these were believed to represent a BMP-independent complication. Excessive back pain was not considered a radicular complication. Following PLF, radiculopathy and wound infection was described.
Methodological Quality Assessment

The methodological quality of all included studies was independently assessed by 2 authors (C.P.H., A.S.H.). Criteria proposed by the Cochrane Back review group were used to assess controlled trials.89 Studies were considered to have a low risk of bias (RoB) if they fulfilled at least 6 of 11 criteria. They were assessed to have a high RoB if they scored 5 or less of the criteria. The quality of noncontrolled studies was determined using a 5-point assessment score.64 We considered them high quality if they fulfilled at least 4 of 5 criteria. A consensus meeting was used to resolve disagreements and a third review author was involved (A.L.).

Statistical Analysis

Event rates are shown as percentages, with a 95% CI. Results of meta-analyses are reported as ORs, with a 95% CI. Due to uncontrollable differences between included studies, such as patient disease severity, definition of fusion, cage and additional instrumentation used, surgeon skill level, and specific surgical technique, it could not be assumed that the true effect of BMP dosing was the same for all studies within their respective groups. We therefore selected a random-effects model, which acknowledges that the effect of treatment may vary from study to study because of heterogeneity. Rather, it assumes that the effect of treatment among studies follows a distribution. The random-effects model was better suited to our study and was used to analyze the fusion and complication data. The Cochrane test (Q-statistic) and the I² statistic were used to assess heterogeneity.39 Publication bias was estimated graphically by assessing funnel plot asymmetry because the low number of included studies did not permit statistical assessment of publication bias.38 (Supplementary Fig. 1). Correlations between fusion rate or complication rate and BMP dose were estimated by calculating the Pearson correlation coefficient. A p value ≤ 0.05 was considered significant. All p values were 2-sided. Statistical analyses were performed using CMA (version 2.2) or SPSS (version 18.0 for Macintosh) software.

Results

Anterior Cervical Discectomy and Fusion

Seven publications5,8,17,34,74,81,87 (Table 1) were included in our analysis. Two studies were prospective5,17 and 5 were retrospective.8,34,74,81,87 Four studies included a control group4,17,34,87 Baskin and colleagues4 provided the only controlled study that carried a low RoB; all other studies were considered high RoB. All noncontrolled studies also exhibited major design weaknesses. Together, we included 397 patients who underwent ACDF. Control patients (98 patients, 24.7%) received either iliac crest bone graft4,17 or autograft shavings placed within a fibular allograft spacer.4,74,87 BMP was used in 299 (75.3%) procedures, with the dose ranging from 0.2 mg to 2.1 mg per level.

Overall, fusion rates exhibited a tendency to be higher in ACDF supplemented with BMP (96.8%, CI 92.2%–98.7%) compared with ACDF lacking BMP (90.8%, CI 82.1%–95.5%, p = 0.068; Fig. 2A and Supplementary Fig. 2A). In constructs devoid of BMP, fusion rates were lower in multilevel ACDFs (Table 2). The fusion rates of ACDF supplemented with 0.2–0.6 mg/level were similar to the rates observed in constructs with 1.1–2.1 mg/level (97.2%, CI 87.4%–99.4% vs 97.7%, CI 79.8%–99.5%). Moreover, even the lowest BMP dose (0.2–0.6 mg/level) achieved high fusion rates in multilevel constructs (2 and 3 levels). Fusion rates of single-level constructs were not significantly altered by the use of BMP (BMP group fusion rate: 97.0%, CI 86.3%–99.4% vs control group fusion rate: 94.3%, CI 76.1%–98.9%). In contrast, application of BMP in multilevel ACDF (two or more levels of fusion) led to a significant improvement (p < 0.001) of the fusion rate (BMP group fusion rate: 98.4%, CI 95.4%–99.4% vs control group fusion rate: 85.8%, CI 77.4%–91.4%). A meta-analysis performed on studies with control groups mirrored these findings.

### Table 1. Overview of ACDF

<table>
<thead>
<tr>
<th>Authors &amp; Year</th>
<th>Study Design</th>
<th>No. of Patients</th>
<th>BMP Dose per Level (mg/level)</th>
<th>% Patients w/ Fusion (n)</th>
<th>% Patients Readmitted, Reintubated, or w/ Tracheostomy (n)</th>
<th>% Patients Returned to OR (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>w/ BMP</td>
<td>Control</td>
<td>BMP</td>
<td>Control</td>
</tr>
<tr>
<td>Frenkel et al., 2013A</td>
<td>Retrospective</td>
<td>9</td>
<td>9</td>
<td>†</td>
<td>0.2–0.4</td>
<td>100 (9)</td>
</tr>
<tr>
<td>Baskin et al., 2003</td>
<td>Prospective</td>
<td>33</td>
<td>18</td>
<td>15</td>
<td>0.6</td>
<td>100 (18)</td>
</tr>
<tr>
<td>Stachnik et al., 2011</td>
<td>Retrospective</td>
<td>30</td>
<td>30</td>
<td>0</td>
<td>0.6</td>
<td>100 (30)</td>
</tr>
<tr>
<td>Frenkel et al., 2013B</td>
<td>Retrospective</td>
<td>7</td>
<td>7</td>
<td>†</td>
<td>0.7–1.1</td>
<td>100 (7)</td>
</tr>
<tr>
<td>Buttermann, 2008</td>
<td>Prospective</td>
<td>66</td>
<td>30</td>
<td>36</td>
<td>0.9</td>
<td>96.7 (29)</td>
</tr>
<tr>
<td>Vaidya et al., 2007</td>
<td>Retrospective</td>
<td>46</td>
<td>22</td>
<td>24</td>
<td>1.0</td>
<td>100 (22)</td>
</tr>
<tr>
<td>Frenkel et al., 2013C</td>
<td>Retrospective</td>
<td>31</td>
<td>8</td>
<td>23</td>
<td>1.4–2.1</td>
<td>100 (8)</td>
</tr>
<tr>
<td>Boakye et al., 2005</td>
<td>Retrospective</td>
<td>24</td>
<td>24</td>
<td>0</td>
<td>2.1</td>
<td>100 (24)</td>
</tr>
<tr>
<td>Shields et al., 2006</td>
<td>Retrospective</td>
<td>151</td>
<td>151</td>
<td>0</td>
<td>2.1</td>
<td></td>
</tr>
</tbody>
</table>

* When a single paper is cited more than once to separate out the results for different BMP doses, the dosage groups are designated by letter labels next to the year of the study (see also Table 5). Only the most severe complications (readmission to the hospital, reintubation and/or tracheostomy, or return to the operating room) that were reliably documented by the included studies were analyzed.

† Control cases are included in Frenkel et al., 2013C.
There was a tendency toward higher fusion rates in constructs using BMP compared with constructs without (OR 2.9, CI 0.797–10.839, p = 0.105; Fig. 2B).

Due to the inconsistent definition of dysphagia in the current literature we avoided an interstudy comparison.

Rather, we focused our analysis on the most severe complications requiring reoperation, readmissions, reintubations, or tracheostomies—all of which were reliably reported in the included studies (Table 1, Fig. 2A). Including only studies with control groups, our meta-analysis revealed a
TABLE 2. Fusion rate for ACDF as a function of levels fused

<table>
<thead>
<tr>
<th>No. of Levels Fused</th>
<th>% in Control Group (n)</th>
<th>% w/ 0.2–0.6 mg BMP/Level (n)</th>
<th>% w/ 0.7–2.1 mg BMP/Level (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>97.06 (34)</td>
<td>100 (10)</td>
<td>100 (108)</td>
</tr>
<tr>
<td>2</td>
<td>88.61 (79)</td>
<td>100 (35)</td>
<td>100 (156)</td>
</tr>
<tr>
<td>3</td>
<td>92.6 (27)</td>
<td>100 (12)</td>
<td>98.88 (89)</td>
</tr>
<tr>
<td>4</td>
<td>62.5 (8)</td>
<td>NA</td>
<td>100 (14)</td>
</tr>
</tbody>
</table>

NA = not available.

2.4-fold increased likelihood of complications when BMP was used in ACDF (Fig. 2C). However, this difference did not reach statistical significance (p = 0.247). In control patients, the rate of these complications was 2.8%, with a CI of 0.8%–9% (Supplementary Fig. 2A). The complication rate was identical in the ACDF cohort without BMP when compared to ACDF in which 0.2–0.6 mg of BMP per level was used. Complication rates were higher in cohorts in which higher BMP doses per level were used. Thus, complication rates were 7.9% (CI 3.2%–18.4%) with 0.7–1.05 mg BMP per level, and 10.2% (CI 6.4%–15.7%) with 1.1–2.1 mg BMP per level. The rate of complications was significantly higher in ACDF supplemented with 1.1–2.1 mg BMP compared with controls (p < 0.05). The rate of complications exhibited a significant positive correlation with the dose of BMP used per level (Pearson correlation coefficient 0.96, p < 0.05).

Posterior Cervical Fusion

We included 6 studies reporting on PCF supplemented with BMP25,29,37,40–42 (Table 3); all 6 were retrospective. Two included a control group,25,39 and both of these studies were associated with a high RoB. Four studies were case series,29,37,41,42 of which all but one25 were of low quality. Our analysis included studies describing posterior cervical arthrodesis only,25,48 but also reports on PCF extending to the occiput29,42 or thoracic spine.29,37,40 A total of 351 patients were included in our analysis, with 248 patients (70.7%) receiving BMP compared with 103 control patients (29.3%) who were not exposed to BMP. The BMP dose used ranged from 1.8 mg/level to 4.5 mg/level.

None of the included studies provided fusion rates of control cohorts. Four studies assessed fusion in PCF supplemented with BMP29,37,41,42 (Fig. 3, Supplementary Fig. 2B). The arthrodesis rate of PCF supplemented with BMP was 95.5%, with a CI of 80.0%–99.1%. Fusion rates were similar in constructs with low (≤ 2.1 mg) or high (> 2.1 mg) doses of BMP per level (95.6% and 95.3%, respectively). Data for meta-analysis on BMP-dependent fusion rates in PCF are currently lacking.

The most frequently reported complications in PCF consisted of wound infections and neurological complications including postoperative C-5 and C-8 palsies (Table 3, Fig. 3, Supplementary Fig. 2B). In control patients the combined complication rate was 10.9%, with a CI of 3.8%–27.5%. Complication rates were 2.9%, with a CI of 0.7%–11.1% in PCF supplemented with ≤ 2.1 mg BMP per level. In patients who received > 2.1 mg of BMP per level, complication rates were similar to those in control patients (9.6%, CI 4%–21.8%). A meta-analysis of the 2 studies with control groups did not suggest a BMP-dependent increase of complications in PCF (Fig. 3B).

Anterior Lumbar Interbody Fusion

A total of 9 publications reported on the effect of BMP on clinical efficacy and complications following ALIF3,14,16,20,22,48,55,77,83 (Table 4). Five studies involved prospective cohorts,14,16,48,77,83 of which 3 contained a control group.16,77,83 Of 6 included case-control studies, 4 had a low RoB16,22,77,83 and 2 had a high RoB.55 All 3 included cases series,5,48 were of high quality. Thus, our analysis included a total of 1007 patients who underwent ALIF with (631 patients, 62.7%) or without (376 patients, 37.3%) BMP. The BMP dose ranged from 2.1 to 12.0 mg per arthrodesis level.

Two studies included in the current analysis assessed fusion rates in ALIF without BMP. They reported a fusion rate of 79.1% with a CI of 57.6%–91.3%. Use of BMP led to a statistically significant increase of fusion rate in ALIF (96.9%, CI 92.3%–98.8%, p < 0.01). Fusion rates were similar among various doses of BMP analyzed (Fig. 4A, Supplementary Fig. 3A). Meta-analysis confirmed a significantly increased likelihood of fusion in constructs supplemented with BMP (OR 22.1, p < 0.001, Fig. 4B).

The most frequent BMP-related complications were retrograde ejaculation, endplate resorption, and graft sub-

TABLE 3. Overview of PCF

<table>
<thead>
<tr>
<th>Authors &amp; Year</th>
<th>Study Design</th>
<th>Control Group</th>
<th>Total</th>
<th>w/ BMP</th>
<th>Control</th>
<th>BMP Dose per Level (mg/level)</th>
<th>% Patients w/ Fusion (n)</th>
<th>% Patients w/ Neurological Complications (n)</th>
<th>% Patients w/ Wound Infection (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hamilton et al., 2011</td>
<td>Retrospective</td>
<td>No</td>
<td>53</td>
<td>53</td>
<td>0</td>
<td>1.8</td>
<td>100 (53)</td>
<td>0 (0)</td>
<td>1.9 (1)</td>
</tr>
<tr>
<td>Hiresmath et al., 2012</td>
<td>Retrospective</td>
<td>Yes</td>
<td>83</td>
<td>16</td>
<td>67</td>
<td>2.0</td>
<td>63 (1)</td>
<td>4.5 (3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Hodges et al., 2009</td>
<td>Retrospective</td>
<td>No</td>
<td>29</td>
<td>29</td>
<td>0</td>
<td>2.1</td>
<td>89.7 (26)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Crawford et al., 2009</td>
<td>Retrospective</td>
<td>Yes</td>
<td>77</td>
<td>41</td>
<td>36</td>
<td>3.6</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>14.6 (6)</td>
</tr>
<tr>
<td>Dorward et al., 2013</td>
<td>Retrospective</td>
<td>No</td>
<td>57</td>
<td>57</td>
<td>0</td>
<td>4.4</td>
<td>89.5 (51)</td>
<td>3.5 (2)</td>
<td>7.02 (4)</td>
</tr>
<tr>
<td>Hood et al., 2014</td>
<td>Retrospective</td>
<td>No</td>
<td>52</td>
<td>52</td>
<td>0</td>
<td>4.5</td>
<td>100 (50)*</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

* Only 50 of 52 patients had sufficient radiographic follow-up; all 50 of those achieved solid fusion.
sidence. When combining reports in which various BMP concentrations were used, there was a trend toward a higher rate of complications with the use of BMP (8.9%, CI 5.3%–14.6%) compared with procedures performed without BMP (3.5%, CI 1.5%–8.35%, p = 0.069; Supplementary Fig. 3A). Use of a low dose of BMP (≤ 4.2 mg/level) increased the rate of complications to 7%, with a CI from 3.4% to 13.9%. The use of 8.4–12.0 mg BMP per level was associated with a 17.7% (CI 4.8%–47.9%) rate of complications. The rate of complications exhibited a significant positive correlation with the dose of BMP used per level (Pearson correlation coefficient 0.98, p < 0.05). Meta-analysis confirmed an increased likelihood of complications in ALIF supplemented with BMP (OR 3.8, p < 0.05; Fig. 4C). Moreover, there was a weak trend toward an association between the likelihood of complications and the dose of BMP used (Pearson correlation coefficient 0.50, p = 0.31).

Transforaminal Lumbar Interbody Fusion

Seventeen articles that reported complication rates in PLIF and/or TLIF were included in our analysis2,23,24,35,38,47,50,52,57,59,60,62,63,65,69,71,75 (Table 5). Three studies were prospective2,47,63 and 360,62,71 were retrospective case-control studies. Three studies were associated with a low RoB.62,63,71 We also included 1 prospective59 and 10 retrospective23,24,35,38,50,52,57,65,69,75 case series. The majority of these studies (7 of 11) were of high quality. The studies comprised a total of 2156 patients, of whom 1852 (85.9%) received rhBMP-2 versus 52 (2.4%) patients who underwent grafting with either local47 or iliac crest autologous bone.2,62 The dose of rhBMP-2 placed into the disc space varied between studies, ranging from 1.4 mg to 12.0 mg. The BMP was placed either into the anterior portion of the prepared disc space (3 studies),2,24,65 into the cage (5 studies),35,50,59,63,71 or both (7 studies).38,47,52,57,60,62,70 Only 2 studies used BMP morselized with bone graft.23,75 In 6 studies BMP (3.0–8.0 mg per level) was also used in the posterolateral gutter.2,23,24,35,38,47 The TLIF procedure was performed in 12 studies2,23,24,35,38,50,52,60,62,63,65,69,75 and PLIF in 2.35,59 Both TLIF and PLIF were included in 3 studies.47,57,71

![PCF - Event rates for Fusion & Complication](image)

**PCF - Event rates for Fusion & Complication**

**PCF - Meta-analysis for Complications**

<table>
<thead>
<tr>
<th>Study name</th>
<th>Statistics for each study</th>
<th>BMP mg/level</th>
<th>Odds ratio</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Z Value</th>
<th>P Value</th>
<th>Odds ratio and 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hiremath et al</td>
<td>0.339</td>
<td>0.041</td>
<td>2.842</td>
<td>-0.997</td>
<td>0.319</td>
<td>2.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crawford et al</td>
<td>6.000</td>
<td>0.105</td>
<td>23.617</td>
<td>0.242</td>
<td>0.809</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Heterogeneity: Q: 1.4, I²: 70.3, (P = 0.71)**

**FIG. 3.** Average fusion and complication rates as a function of BMP dose used per level in PCF (A). Meta-analysis for complications in PCF supplemented with BMP (B).
The fusion rate of TLIF performed without BMP was 93.0%, with a CI from 78.1% to 98.0% (Fig. 5, Supplementary Fig. 3B). The fusion rate in all constructs supplemented with BMP was 95.0% (CI 92.8%–96.5%). There was a weak trend toward slightly higher fusion rates in constructs supplemented with BMP; however, this difference was not statistically significant. Subgroup analysis revealed that the use of larger BMP doses per disc space did not translate into increased fusion rates. Thus, the fusion rates of constructs supplemented with the lowest dose (≤ 4 mg/level) were similar to constructs supplemented with the highest doses (8.5–12.0 mg/level; 95.5% and 94.3%, respectively). Our exploratory meta-analysis was in line with these findings. There was a nonsignificant trend toward a higher likelihood of fusion in TLIFs supplemented with BMP (OR 4.5, p = 0.098; Fig. 5B) that was independent of dose.

Our analysis focused on complications most frequently described in the included studies: radiculopathy, heterotopic bone formation, wound infections, and seromas. These complications occurred at a rate of 5.1% (CI 1.3%–18.3%) in control patients (Supplementary Fig. 3B). There was a tendency toward a higher rate of complications in TLIF supplemented with BMP (7.1%, CI 4.8%–10.5%). Subgroup analysis revealed that the rate of complications was not positively correlated with the dose of BMP used by level. Rather, the rate of complications was twice as high in constructs supplemented with low doses of BMP (≤ 4 mg/level), but compared with control constructs this is not statistically significant. Interestingly, complication rates in patients who received larger doses of BMP were similar to those in control patients. An exploratory meta-analysis confirmed a weak trend toward a higher likelihood of complications in TLIFs supplemented with BMP (Fig. 5C).

### Posterolateral Lumbar Fusion

Our analysis included 9 papers26–28,36,45,53,76,82,86 that reported the efficacy and complications of BMP in PLF procedures (Table 6). Seven studies were case-control studies,26–28,36,45,53,76,82 whereas 2 were case series.36,82 Seven studies were prospective26–28,45,76,82,86 and had a low RoB. 36,53 The 2 retrospective studies had a high RoB.36,53 Thus, our analysis included a total of 1978 patients; 1552 (78.5%) with BMP versus 426 (21.5%) without BMP. The BMP dose used per level ranged from 4.2 to 42.0 mg. In the control group, 75.3% (CI 64.1%–84.0%) of patients achieved successful arthrodesis (Fig. 6, Supplementary Fig. 4). PLF supplemented with BMP showed a trend toward a higher fusion rate (88.5%, CI 78.6%–94.2%) compared with controls (p = 0.066). There was weak positive correlation between fusion rate and BMP dose (Pearson correlation coefficient 0.678, p = 0.32). When only considering constructs supplemented with ≥ 8.5 mg BMP per level, the fusion rate was 95.2% (CI 90.1%–97.8%); significantly greater than in the control group (p < 0.001). Meta-analysis confirmed a higher likelihood of arthrodesis of PLF supplemented with BMP compared with controls (p < 0.01, Fig. 6B).

The most common complications following PLF were wound infections and radiculopathy (Supplementary Fig. 4). These complications occurred in 9.6% (CI 4.8%–
18.4%) of control patients. The rate was similar in patients who received BMP (8.2%, CI 4.1%–16.0%). Moreover, the BMP dose had no significant impact on the complication rate. Thus, the rate of complications in PLF supplemented with the highest BMP doses (35–40 mg/level, 14.5%; CI 6.1%–30.8%) was similar to the rate in control patients. Meta-analysis confirmed this observation (Fig. 6C).

**Discussion**

Autologous iliac bone graft remains the gold standard for promoting fusion in spinal arthrodesis procedures. However, the high donor site morbidity, which has been reported to affect up to 47%–60% of patients, has fueled an intense search for bone graft substitutes and enhancers.
<table>
<thead>
<tr>
<th>Authors &amp; Year</th>
<th>Study Design</th>
<th>Ctl Group</th>
<th>No. of Pts w/ BMP (mg)</th>
<th>Total per Level (mg/level)</th>
<th>Int (mg)</th>
<th>Post (mg)</th>
<th>BMP Dose</th>
<th>% Pts w/ Fusion (n)</th>
<th>% Pts w/ Radiculopathy</th>
<th>% Pts w/ Het/Ect Bone Formation (n)</th>
<th>% Pts w/ Seroma (n)</th>
<th>% Pts w/ Wound Infection (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mannion et al., 2011</td>
<td>Retro</td>
<td>No</td>
<td>30</td>
<td>30</td>
<td>0</td>
<td>1.4</td>
<td>1.4</td>
<td>0</td>
<td>96.7 (29)</td>
<td>3.3 (1)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Crandall et al., 2013A</td>
<td>Retro</td>
<td>No</td>
<td>69</td>
<td>69</td>
<td>0</td>
<td>2.0−4.0</td>
<td>2.0−4.0</td>
<td>0</td>
<td>98.6 (68)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Geibel et al., 2009A</td>
<td>Retro</td>
<td>No</td>
<td>11</td>
<td>11</td>
<td>0</td>
<td>6.0</td>
<td>3.0</td>
<td>3.0</td>
<td>100 (11)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Roh et al., 2013</td>
<td>Retro</td>
<td>Yes</td>
<td>321</td>
<td>95</td>
<td>0†</td>
<td>3.07</td>
<td>3.07</td>
<td>0</td>
<td>90.1 (86)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Owens et al., 2011</td>
<td>Retro</td>
<td>No</td>
<td>204</td>
<td>204</td>
<td>0</td>
<td>12.0</td>
<td>4.0</td>
<td>8.0</td>
<td>5.9 (12)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>2.0 (4)</td>
</tr>
<tr>
<td>Kuklo et al., 2004</td>
<td>Retro</td>
<td>No</td>
<td>22</td>
<td>22</td>
<td>0</td>
<td>4.2</td>
<td>4.2</td>
<td>0</td>
<td>95.5 (21)</td>
<td>4.5 (1)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Joseph &amp; Rampersaud, 2007</td>
<td>Pro</td>
<td>Yes</td>
<td>33</td>
<td>23</td>
<td>10</td>
<td>4.2</td>
<td>4.2</td>
<td>0</td>
<td>100 (23)</td>
<td>90.0 (9)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Nandyala et al., 2014</td>
<td>Pro</td>
<td>Yes</td>
<td>52</td>
<td>26</td>
<td>0†</td>
<td>4.2</td>
<td>4.2</td>
<td>0</td>
<td>92.3 (24)</td>
<td>3.8 (1)</td>
<td>3.8 (1)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Mindea et al., 2009A</td>
<td>Retro</td>
<td>No</td>
<td>43</td>
<td>35</td>
<td>8</td>
<td>4.2</td>
<td>4.2</td>
<td>0</td>
<td>11.4 (4)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Singh et al., 2009A</td>
<td>Retro</td>
<td>No</td>
<td>344</td>
<td>344</td>
<td>0</td>
<td>4.2</td>
<td>4.2</td>
<td>0</td>
<td>95.1 (327)</td>
<td>2.6 (9)</td>
<td>2.6 (9)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Rihn et al., 2009A</td>
<td>Retro</td>
<td>No</td>
<td>48</td>
<td>48</td>
<td>0</td>
<td>8.4</td>
<td>4.2</td>
<td>4.2</td>
<td>95.8 (46)</td>
<td>14.6 (7)</td>
<td>2.1 (1)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Geibel et al., 2009B</td>
<td>Retro</td>
<td>No</td>
<td>37</td>
<td>37</td>
<td>0</td>
<td>8.4</td>
<td>4.2</td>
<td>4.2</td>
<td>100 (37)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Anand et al., 2006</td>
<td>Pro</td>
<td>Yes</td>
<td>100</td>
<td>85</td>
<td>15</td>
<td>8.4</td>
<td>4.2</td>
<td>4.2</td>
<td>100 (85)</td>
<td>93.3 (14)</td>
<td>3.5 (3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Knox et al., 2011</td>
<td>Retro</td>
<td>No</td>
<td>58</td>
<td>58</td>
<td>0</td>
<td>5.0</td>
<td>5.0</td>
<td>0</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Helgeson et al., 2011</td>
<td>Retro</td>
<td>No</td>
<td>23</td>
<td>23</td>
<td>0</td>
<td>12.0</td>
<td>6.0</td>
<td>6.0</td>
<td>82.6 (19)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Corenman et al., 2013</td>
<td>Retro</td>
<td>No</td>
<td>36</td>
<td>36</td>
<td>0</td>
<td>12.0</td>
<td>6.0</td>
<td>6.0</td>
<td>100 (36)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

(continued)
In 2002 the US FDA approved BMP for single-level ALIF (levels L2–S1) for treatment of degenerative disc disease, Grade I spondylolisthesis, and/or retrolisthesis with the lumbar tapered fusion device LT-CAGE (Medtronic Sofamor Danek). Depending on the size of the LT-CAGE, the FDA recommended between 4.2 mg and 12 mg BMP per level.

Subsequently, a steep rise in off-label BMP applications ensued in various spinal arthrodesis procedures. Given the lack of guidelines or recommendations, wide ranges of the BMP doses are being used. This is reflected in the studies included in the current report. Doses of BMP used for specific procedures varied by a factor of 7.5 (range 2.5 [PCF] to 10.5 [ACDF]) between the studies that used the smallest and those that used the largest doses. This would translate into giving arbitrarily either 40 mg or 300 mg of Lasix for the same indication. Serious complications have been reported due to BMP, in particular the following: airway compromise and dysphagia following ACDF, ectopic bone formation with TLIF, and retrograde ejaculation in ALIF. However, the potential impact of BMP dose on the complication rate is unclear.

Anterior Cervical Discectomy and Fusion

Fusion rates in 1-level ACDF are largely unaffected by the use of BMP, but when comparing fusion rates in multilevel (2, 3, and 4 levels) operations, there appears to be a significant (p < 0.001) increase in the rate of fusion. Our data suggest that there is a significantly increased risk of complications with BMP in ACDF (p < 0.05) and that the risk is dose dependent. Perri and colleagues were one of the first groups to describe a case of cervical swelling following a 2-level ACDF, in which the patient required intubation and intensive care. This report lacked information regarding the BMP dose. It has been proposed that an early BMP-mediated inflammatory response exacerbates prevertebral soft-tissue swelling. Vaidya and colleagues demonstrated that the use of BMP leads to significantly greater prevertebral soft-tissue swelling peaking at 1 week after surgery, as determined on lateral cervical radiographs. Cervical swelling associated with use of rh-BMP-2 was reported in 6 of the included studies. Three studies in which the highest dose of BMP (2.1 mg/level) was used report dysphagia rates ranging from 8.3% to 80%. Shields and colleagues reported 13 cervical swelling events (8.6%) leading to breathing or swallowing difficulties. Five of these patients had a prolonged hospital stay and 8 required readmission. Frenkel and colleagues detected dysphagia or cervical swelling in 4 of 5 patients (80%) who received 2.1 mg BMP per level and in 1 of 3 (33.3%) who received 1.4 mg/ml. The latter patient required an emergency room visit due to the severity of his symptoms.

Studies that used the lowest doses of BMP (≤ 0.6 mg per level) reported a dysphagia rate ranging from 0% to 6.7%. Stachniak and colleagues included a very thorough evaluation of postoperative dysphagia using the SWAL-QOL scale. At 6 months after ACDF performed with 0.6 mg BMP per level, 6.7% of patients complained about food frequently sticking in the throat. However, none of the patients complained about dysphagia by 1 year post-
operatively. Importantly, none of the patients required re-admission or reoperation. Similarly, no cases of dysphagia were reported in the other 2 studies in which 0.6 mg BMP per level was used.\textsuperscript{4,34} Lu and colleagues performed a large retrospective study on BMP-related postoperative dysphagia.\textsuperscript{56} The authors used the 5-point SWAL-QOL scale\textsuperscript{58} to grade postoperative dysphagia in patients who underwent ACDF with (BMP dose/level ranging from 0.7 to 2.1 mg, \textit{n} = 100) or without (\textit{n} = 50) BMP. Although the incidence of dysphagia was similar in both groups, the severity of dysphagia was significantly higher in the BMP cohort. The authors found a weak correlation between BMP dose and dysphagia score. Also, the incidence of 4% of patients with percutaneous endoscopic gastrostomy placement is inconsistent with the current literature.

Given the inconsistent definition of neck swelling and dysphagia in the current literature, our analysis included only the most severe complications requiring reoperation, readmissions, reintubations, or tracheostomies. In line with the aforementioned studies, we detected a strong positive correlation between complications and BMP dose. The lowest concentration analyzed (0.2–0.6 mg/level) had

---

**FIG. 5.** Average fusion and complication rates as a function of BMP dose used per level in TLIF (A). There is a tendency toward higher fusion rates in TLIF supplemented with BMP. Meta-analysis for fusion (B) and complication (C) rates in TLIF supplemented with BMP. Studies are listed in order of increasing doses of BMP per level.
a complication rate comparable to the control cohort and a fusion rate similar to the cohort that received the highest dose (1.1–2.1 mg/level). Thus, the currently available data support the use of 0.2–0.6 mg/level of BMP in ACDF. However, future studies are necessary to define the minimum effective BMP dose in ACDF.

**Posterior Cervical Fusion**

The current literature provides no data on fusion rate of the control patient cohort following PCF. In the literature, the rate of cervical spine arthrodesis without BMP ranges between 94.2% and 100%,31,44,51,73 Our estimated fusion rate in PCF supplemented with BMP was only marginally better (95.5%, CI 80.0%–99.1%). A recent paper by Crawford and colleagues suggested a trend toward higher rates of wound complications in PCF supplemented with BMP.25 Thus, these authors reported a 14.6% rate of wound complications in cases of PCF supplemented with 3.6 mg of BMP per level compared with the control cohort (2.8%). Our meta-analysis did not suggest a BMP dose–dependent increase of the complication rate. However, the quality of most included studies was poor. Because there was no difference in fusion rate between low (≤ 2.1 mg) or high (> 2.1 mg) doses of BMP, we would recommend limiting the BMP dose to < 2.1 mg per level. There is a great need for more studies confirming the efficacy of BMP on fusion rates in posterior cervical spine surgery.

**Anterior Lumbar Interbody Fusion**

There was a large number of high-quality studies documenting the effectiveness of BMP in ALIF. Taken together, they documented a highly significant enhancement of arthrodesis in our meta-analysis. Our meta-analysis suggests that complication rates following ALIF supplemented with BMP vary significantly in the reported literature. A funnel plot analysis suggests that there might be publication bias toward increased ALIF complication rates in lower-quality studies (defined as studies with a high SEM, Supplementary Fig. 1). However, this finding was weak, subjective, and the number of included studies was low (n = 6). Rather, we propose that heterogeneity is caused by the various BMP concentrations used in these reports. Meta-analysis revealed a trend toward an association between the likelihood of complications and the dose of BMP used (Pearson correlation coefficient 0.50, p = 0.31). This association was strengthened when also including studies without control groups. The association between complication rate and BMP exhibited a significant positive correlation (Pearson correlation coefficient 0.98, p < 0.05), thus indicating that the heterogeneity was at least partially due to various BMP concentrations used in these reports.

In reports of ALIF supplemented with high doses of BMP (4.3–12.0 mg/level), the rates of endplate resorption and graft subsidence were high. Retrograde ejaculation was assessed in the majority of studies. All but 1 study83 used questionnaires to inquire about quality of erections, quality of ejaculate, and clarity of postejaculation urine to assess retrograde ejaculation. In a retrospective case series, Carragee and colleagues reported 5 cases of retrograde ejaculation in the BMP group (7.2%) compared with 1 case (0.6%) in the control group.20 This patient cohort

---

### Table 6: Overview of PLF

<table>
<thead>
<tr>
<th>Authors &amp; Year</th>
<th>Study Design</th>
<th>Group</th>
<th>No. of Pts</th>
<th>BMP Dose per Level (mg/level)</th>
<th>% Pts w/ Fusion</th>
<th>% Pts w/ Radiculopathy</th>
<th>% Pts w/ Wound Infection</th>
<th>Ctl</th>
<th>BMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee et al., 2010</td>
<td>Retrospective</td>
<td>Yes</td>
<td>75</td>
<td>4.2</td>
<td>62.4 (28)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>2.9 (1)</td>
<td>2.4 (1)</td>
</tr>
<tr>
<td>Delawi et al., 2010</td>
<td>Prospective</td>
<td>Yes</td>
<td>36</td>
<td>18</td>
<td>7.0</td>
<td>62.5 (10)†</td>
<td>5.6 (1)</td>
<td>11.2 (4)</td>
<td>5.6 (1)</td>
</tr>
<tr>
<td>Vaccaro et al., 2008</td>
<td>Prospective</td>
<td>No</td>
<td>36</td>
<td>24</td>
<td>7.0</td>
<td>68.8 (11)§</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>16.7 (4)</td>
</tr>
<tr>
<td>Stambough et al., 2010</td>
<td>Prospective</td>
<td>No</td>
<td>103</td>
<td>46</td>
<td>7.0</td>
<td>97.2 (35) †</td>
<td>0.7 (7)</td>
<td>0 (0)</td>
<td>3.9 (46)</td>
</tr>
<tr>
<td>Glassman et al., 2011</td>
<td>Retrospective</td>
<td>No</td>
<td>1037</td>
<td>46</td>
<td>7.0</td>
<td>97.4 (18)§§</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>4.8 (1)</td>
</tr>
<tr>
<td>Dawson et al., 2009</td>
<td>Prospective</td>
<td>Yes</td>
<td>46</td>
<td>25</td>
<td>12.0</td>
<td>94.7 (186) ††</td>
<td>70.0 (14)</td>
<td>0 (0)</td>
<td>0.7 (7)</td>
</tr>
<tr>
<td>Singh et al., 2006</td>
<td>Prospective</td>
<td>Yes</td>
<td>52</td>
<td>41</td>
<td>12.0</td>
<td>95.9 (180)‡‡</td>
<td>63.6 (7)</td>
<td>0 (0)</td>
<td>0.7 (7)</td>
</tr>
<tr>
<td>Dimar et al., 2009</td>
<td>Prospective</td>
<td>Yes</td>
<td>463</td>
<td>239</td>
<td>40.0</td>
<td>95.9 (180)‡‡</td>
<td>63.6 (7)</td>
<td>0 (0)</td>
<td>0.7 (7)</td>
</tr>
<tr>
<td>Hurlbert et al., 2013</td>
<td>Prospective</td>
<td>Yes</td>
<td>197</td>
<td>98</td>
<td>31.5–42.0</td>
<td>93.5 (186) ††</td>
<td>69.3 (151)‡‡</td>
<td>1.7 (4)</td>
<td>2.2 (5)</td>
</tr>
</tbody>
</table>

*16 patients assessed; †15 patients assessed; ‡6 patients assessed; §19 patients assessed; ¶20 patients assessed; **39 patients assessed; ††169 patients assessed; †‡169 patients assessed; §§93 patients assessed; ¶¶88 patients assessed.
was also included in a retrospective series by Comer and colleagues. All together, 239 patients who received BMP and 233 control patients were analyzed for the incidence of retrograde ejaculation, with a minimum follow-up time of 1 year. Gathering specific data regarding retrograde ejaculation during follow-up office visits as well as with specific questionnaires, the authors determined a retrograde ejaculation rate of 6.3% in patients who underwent ALIF supplemented with 4.2 mg of BMP compared with 0.9% in control patients without BMP exposure.

![Graph showing event rates for Fusion & Complication](image)

**FIG. 6.** Average fusion and complication rates as a function of BMP dose used per level in PLF (A). Constructs supplemented with 8.5 mg BMP per level show a higher fusion rate compared with controls. Meta-analysis for fusion (B) and complication (C) rates in PLF supplemented with BMP. Studies are listed in order of increasing doses of BMP per level.
For our analysis, patients from the earlier report by Carragee and colleagues were excluded,$^{20}$ resulting in a 5.9% rate of retrograde ejaculation in the 170 remaining patients treated with BMP and a 1.7% rate in remaining 59 control patients. Lindley et al. assessed the occurrence of retrograde ejaculation via retrospective review of medical records as well as via a phone interview (in 78.9% of patients) to assess possible indicators of this disorder. The postoperative retrograde ejaculation rate was similar in patients who underwent ALIF with (7.4%) or without (9.8%) BMP. Five of the studies included in our current study did not report the occurrence of retrograde ejaculation.$^{5,14,16,48,77}$

As mentioned before, there is only 1 study to date that has used an objective quantitative method for retrograde ejaculation assessment following ALIF.$^{93}$ Preoperatively, patients underwent a semen and postejaculatory urine analysis that was repeated 3–6 months after ALIF. Tepner and colleagues determined the diagnosis based on laboratory data in 2 patients who had undergone ALIF with BMP (9.5%) and in 2 patients in the control group (10.0%). The retrograde ejaculation rate based on the questionnaire was remarkably different. Thus, 15 patients were identified with possible retrograde ejaculation. The authors concluded that the reliability of a questionnaire for retrograde ejaculation is poor, with a positive predictive value of 26.7%. Because the vast majority of our data regarding BMP-related retrograde ejaculation were based on this unreliable test, a confirmatory study using an objective quantitative method is needed. In conclusion, BMP increased the fusion rate in ALIF, but doses > 4.2 mg/level should be avoided due to the increased risk of endplate resorption and/or graft subsidence.

**Transforminal Lumbar Interbody Fusion**

Our meta-analysis revealed that the use of BMP had only a minimal impact on fusion rate in TLIF. Moreover, complication rates were twice as high using even the lowest dose of BMP (≤ 4.2 mg/level). Radiculopathy was one of the most frequently reported complications with the use of BMP in TLIF. Neurological symptoms corresponded to the dermatomal distribution of the TLIF window and occurred more than 10 days after the procedure in a patient who had previously been asymptomatic postoperatively. This complication was described in 2.3% of those who received BMP/PLIF with BMP in the intervertebral space (43 of 1852 patients), whereas it was not described in patients without BMP. In approximately half of the patients (46.5%), imaging did not reveal a structural cause. One-third of patients (34.9%) who suffered from BMP-related radiculopathy had foraminal heterotopic bone formation. Seromas and hematomas where less frequently found to cause radiculopathy (16.3% and 2.3%, respectively).

Interestingly, the time of symptom onset appears to depend on the specific underlying pathology. Radiculopathy due to seromas was described as early as 10 days post-surgery. Radiculopathy without obvious structural cause was seen 2–4 weeks after surgery. Ectopic foraminal bone formation led to radiculopathy in a delayed fashion (> 2 months). Studies including patients who received between 1.4 and 4.2 mg BMP per disc space reported BMP-related radiculopathy at a rate of 3.7%. Briefly, Mannion et al. reported 1 case (3.3%) of radiculopathy in a patient who developed recurrent L-4 radicular pain 4 weeks postsurgery.$^{37}$ MRI revealed a large inflammatory cyst in conjunction with cage repulsion into the foramen, compressing the L-4 nerve root (during the revision surgery it was noted that the cage did not exert pressure onto the nerve root). Nandyala and colleagues reported 1 case of symptomatic neuroforaminal bone growth in a patient who received 4.2 mg of BMP per disc space.$^{63}$ The patient was treated with foraminotomy and neurolysis and a contralateral minimally invasive TLIF. Mindea and colleagues reported 4 cases of radiculopathy in 35 patients (11%) in their BMP cohort.$^{60}$

In contrast to the aforementioned studies, CT imaging revealed no structural cause for the radiculopathy, and all cases of radiculopathy resolved with medical therapy (pregabalin or gabapentin) within 6 weeks. Singh et al. reported a 3.1% rate of symptomatic neuroforaminal ectopic bone growth.$^{73}$ Nine of these patients had received a small INFUSE kit (4.2 mg) and 1 had received a large INFUSE kit (12.0 mg) per disc space. Owens and colleagues performed TLIF procedures in 204 patients, with a BMP dose of 4.0 mg per intervertebral space.$^{56}$ In their series, BMP-related radiculopathy was caused by symptomatic vertebral osteolysis with foraminal collapse (1 patient), foraminal hematoma (1 patient), or foraminal seroma (4 patients). Six patients complained of persistent radiculopathy without clear etiology on imaging studies. Rihn and colleagues found 7 cases of transient postoperative radiculitis$^{69}$ that was caused by ectopic bone formation in 1 patient and by a deep wound infection in another patient. Anand and colleagues reported on 3 patients (3.5%) with BMP-related transient postoperative radicular pain.$^{2}$

In the cohort of studies applying 4.3–8.4 mg BMP per intervertebral space, 5 cases of radiculopathy were reported (1.1%). Crandall and colleagues reported 5 cases of radiculopathy in patients who received 6.0–8.0 mg of BMP per level.$^{24}$ Three of these cases were caused by ectopic bone formation and 2 by seromas. In patients who received between 8.4 and 12.0 mg BMP per intervertebral space, 1 case of radiculopathy was reported (0.3%). Singh and colleagues described 1 patient who had received a large INFUSE kit (12.0 mg) and developed radiculopathy.$^{73}$ Although the BMP dose used did not correlate with the rate of radiculopathy, the preparation of BMP in the disc space might be of importance. Thus, the rate of symptomatic ectopic bone formation was approximately 3 times higher if BMP was mixed with bone graft rather than either placed anterior to the disc space or within the spacer. In conclusion, the addition of BMP to a TLIF resulted in a minimal increase of fusion rate while increasing complication rates. Future studies should aim to define the minimum effective BMP dose (≤ 4.2 mg/level) to achieve the best benefit-to-complication profile.

**Posterolateral Lumbar Fusion**

The majority of studies on PLF were of good quality. Thus, we can conclude that substitution of PLF with 8.5–12.0 mg/level of BMP results in higher fusion rates. Moreover, complication rates appear to remain comparable to those in control patients regardless of the dose of BMP.
used. Thus, 8.5–12.0 mg of BMP per level appears to be a safe and effective dose for PLF.

Limitations of the Analysis

The quality of the results of a meta-analysis rely heavily on the quality of included studies. We attempted to be entirely compliant with the PRISMA guidelines, and we admit that therefore the quality of our conclusions differs among various procedures analyzed. Thus, the average quality of studies available for ACDF and PCF was poor, and therefore further studies are warranted. A body of high-quality literature is available for both ALIF and PLF, which allows for more definitive conclusions. We admit that given our stringent inclusion criteria we were only able to include a fraction of the available literature. Several high-quality studies were not included because they did not provide sufficient information regarding the dose of BMP used per level in relation to outcome and complications. We admit that there are a multitude of other factors that may influence the efficacy of bony union and the complication rates. This includes for example patient selection, comorbidities, surgical technique, the preparation of BMP, the use of intervertebral cages, and supplemental instrumentation. Although we performed a random-effects analysis on our data to account for these variables, our results were certainly biased by these factors. Thus, the results of this meta-analysis should serve as the foundation for future well-designed studies to provide better guidelines on the use and dosing of BMP in spinal arthrodesis surgery.

Conclusions

Our meta-analysis suggests that there is a tremendous disparity in the BMP dosing used for specific spinal fusion procedures. The efficacy and complications associated with BMP are highly dependent on the type of procedure. Particularly when BMP is applied into the intervertebral space the therapeutic window appears to be narrow, and therefore future studies are needed to further investigate these doses. On the other hand, when used in the posterolateral gutter of the lumbar spine, the efficacy appears great and complication rates acceptable. The quality of studies was too poor to expand this statement to the posterior cervical spine, where high-quality studies are lacking. In summary, the range of BMP doses used for various spinal arthrodesis procedures is great, and our study should provide some guidance for the spine surgeon in need of bone graft substitutes. It should also stimulate new studies to further define safe and efficacious BMP dosing for various applications.

References


posteriorlateral lumbar spine fusion in elderly patients. *Eur Spine J* 19:924–930, 2010


87. Vaidya R, Carp J, Sethi A, Bartol S, Craig J, Les CM: Complications of anterior cervical disectomy and fusion using...
recombinant human bone morphogenetic protein-2. Eur Spine J 16:1257–1265, 2007

Disclosures
Dr. Levi receives educational honoraria from DePuy Synthes and Medtronic. Dr. Hofstetter performs consulting and teaching for Globus Medical and InVivo therapeutics.

Author Contributions
Conception and design: all authors. Acquisition of data: all authors. Analysis and interpretation of data: all authors. Drafting the article: all authors. Critically revising the article: all authors. Reviewed submitted version of manuscript: Levi. Approved the final version of the manuscript on behalf of all authors: Levi. Statistical analysis: Levi, Hofstetter.

Supplemental Information
Online-Only Content
Supplemental material is available with the online version of the article.

Correspondence
Allan D. Levi, Department of Neurological Surgery, University of Miami Miller School of Medicine, Lois Pope Life Center, 1095 N.W. 14th Terrace (D4-6), Miami, FL 33136. email: alevi@med.miami.edu.