

# Change in Off-Label Use of Bone Morphogenetic Protein in Spine Surgery and Associations with Adverse Outcome

Jashvant Poeran<sup>1</sup> Mathias Opperer<sup>2</sup> Rehana Rasul<sup>1</sup> Madhu Mazumdar<sup>1</sup> Federico P. Girardi<sup>3</sup>  
Alexander P. Hughes<sup>3</sup> Stavros G. Memtsoudis<sup>2</sup> Vassilios Vougioukas<sup>4</sup>

<sup>1</sup>Department of Population Health Science and Policy, Institute for Healthcare Delivery Science, Icahn School of Medicine at Mount Sinai, New York, New York, United States

<sup>2</sup>Department of Anesthesiology, Hospital for Special Surgery, New York, New York, United States

<sup>3</sup>Department of Orthopedic Surgery, Hospital for Special Surgery, New York, New York, United States

<sup>4</sup>Department of Neurosurgery, Albert Ludwig University of Freiburg, Freiburg, Germany

**Address for correspondence** Jashvant Poeran, MD, PhD, Department of Population Health Science and Policy, Institute for Healthcare Delivery Science, Icahn School of Medicine at Mount Sinai, 1425 Madison Avenue (Box 1077), New York, NY 10029, United States (e-mail: jashvant.poran@mountsinai.org).

Global Spine J 2016;6:650–659.

## Abstract

**Study Design** Retrospective cohort study.

**Objective** The U.S. Food and Drug Administration issued a warning in 2008 against off-label bone morphogenetic protein (BMP-2) use. We aimed to determine (off-label) BMP-2 use in two periods and associations with complications.

**Methods** We included 340,393 patients undergoing spinal fusions from the Premier Perspective database (2006 to 2012). BMP-2 use was determined from billing in 2006 to 2008 versus 2009 to 2012. Outcomes included revisions, length of hospital stay (LOHS), and cost of hospital stay (COH). Multilevel regressions measured associations between BMP-2 and outcomes; odds ratios (ORs) and 95% confidence intervals (CIs) are reported.

**Results** BMP-2 use decreased from 18.7% in 2006 to 11.5% in 2012. Off-label use remains but is decreasing, particularly for cervical anterior (5.1 versus 2.0%) and cervical posterior procedures (15.3 versus 8.5%; both  $p < 0.01$  comparing 2006 to 2008 with 2009 to 2012). BMP-2 remains associated with increased LOHS (median 2 versus 3 days; both periods) and COH (median \$15,455 versus \$27,881 in 2006 to 2008; \$17,007 versus \$30,331 in 2009 to 2012). Adjusted ORs for the association between BMP-2 and adverse outcomes were generally lower in 2009 to 2012 compared with 2006 to 2008. Most notably, we demonstrate lower ORs for revision after cervical fusions in 2009 to 2012 (OR 1.67, CI 1.01 to 2.78) compared with 2006 to 2008 (OR 2.43, CI 1.66 to 3.54).

**Conclusions** Using a previously untapped data source, we show decreased (off-label) BMP-2 use in spinal fusions, particularly in cervical fusions. Although there was a tendency of decreased odds in 2009 to 2012, higher resource utilization and odds for complications remain in patients using BMP-2. A national registry or prospective observational studies will benefit the ongoing discussion.

## Keywords

- bone morphogenetic protein
- off-label
- complications
- resource utilization
- time trends
- spinal fusion

received  
June 1, 2015  
accepted after revision  
November 24, 2015  
published online  
January 15, 2016

DOI <http://dx.doi.org/10.1055/s-0036-1571284>.  
ISSN 2192-5682.

© 2016 Georg Thieme Verlag KG  
Stuttgart · New York

License terms



## Introduction

Recombinant human bone morphogenetic protein (BMP, specifically BMP-2), a growth factor capable of inducing the formation of new bone, is used in a significant part of the >400,000 spinal fusion procedures in the United States as an alternative to iliac crest bone graft (autograft).<sup>1,2</sup> The U.S. Food and Drug Administration (FDA) approved the use of BMP-2 (InFuse, Medtronic, Memphis, Tennessee, United States) in 2002 for anterior interbody lumbar fusions.<sup>3</sup>

Since then, its use has been the subject of numerous small studies and several large-scale observational studies, fueling the debate on its safety and increased costs of hospitalization.<sup>1,4-9</sup> Specifically, one study showed 85% of spinal fusions using BMP-2 were for off-label applications.<sup>8</sup> The most serious complications following BMP-2 use have been reported for anterior cervical spine surgery (e.g., airway compression), which resulted in the FDA issuing an official warning in 2008.<sup>10</sup>

Despite the multitude of studies, many small studies addressing off-label use and safety of BMP-2 are burdened by authors' conflicts of interest.<sup>6,11</sup> Moreover, large-scale studies have not used the most recent data,<sup>1,7-9</sup> only included a subset of spinal fusions,<sup>7,9</sup> or did not account for hospital variation.<sup>1,7,9</sup> Furthermore, the main effect of interest, BMP-2 use, was defined by International Classification of Diseases, 9th revision (ICD-9) codes that do not differentiate between BMP-2 and BMP-7 (Stryker Biotech, Hopkinton, Massachusetts, United States).<sup>1,7-9</sup> The latter is less often used and was approved by the FDA in 2001 for use as an alternative to recalcitrant long bone nonunions under a so-called "humanitarian device exemption."<sup>12</sup> Using a nationwide claims-based data set previously untapped for these studies while also containing more recent data, we therefore sought to determine (1) off-label use of BMP-2 and changes between two periods divided by the year of the FDA warning and (2) associations of its use with adverse outcome in the two periods while accounting for hospital-level effects (i.e., patients correlated to each other in the same hospital and variation across hospitals) by multilevel modeling.

## Methods

### Data Source and Study Design

For this retrospective cohort study, we used the Premier Perspective database (Premier Inc., Charlotte, North Carolina, United States) containing hospital discharges for the period January 2006 to October 2012 for ~20 to 25% of U.S. hospitals.<sup>13</sup> The Premier Perspective database provides complete information on billing from a patient's hospitalization as well as information regarding ICD-9 Clinical Modification codes and Current Procedural Terminology codes. Standardized codes, such as ICD-9 and Current Procedural Terminology codes, are ascertained to be valid for the time period the patient record is reported. Demonstrating its potential, the Premier Perspective data set is used by an increasing number of study groups addressing a variety of clinical questions.<sup>14</sup> As the data set meets the requirements of deidentification as defined by the Health Insurance Portability and Accountability Act, it was exempt from consent requirements of our Institutional Review Board.

### Study Sample

Following previous studies, we defined our cohort based on ICD-9 procedure codes.<sup>1,15</sup> We included all patients undergoing elective primary spinal fusions: cervical (ICD-9 codes 81.02, 81.03), thoracic (ICD-9 codes 81.04, 81.05), or lumbar (ICD-9 codes 81.06 to 81.08). Cases with unknown gender ( $n = 6$ ), multiple procedures ( $n = 24,448$ ), or unknown discharge status ( $n = 879$ ) were excluded. Because our main effect of interest was the use of BMP-2, we also excluded patients undergoing spinal fusions in which other or unspecified BMPs were used ( $n = 4,015$ ).

### Main Effect of Interest

Our main effect of interest was the use of BMP-2, which was extracted from billing data because ICD-9 codes do not differentiate between BMP-2 and BMP-7. The main assumption is that BMP-2 use that was actually billed for provides an accurate estimate of its use.

### Additional Study Variables

Patient demographics included age, gender, race (white, black, Hispanic, other), and smoking status (ICD-9 codes 305.1X, V15.82). Health care-related variables included type of insurance (commercial, Medicaid, Medicare, uninsured, other), hospital location (rural, urban), hospital size (<300, 300 to 499, ≥500 beds), hospital teaching status, and the mean annual number of spinal fusions per hospital. Procedure-related variables were derived from ICD-9 codes as used in previous studies and included the diagnosis associated with the fusion (degenerative disease, spondylolisthesis, deformity, spinal stenosis, other diagnoses), the segment/approach of fusion, number of vertebral levels fused, use of an interbody device, and year of procedure.<sup>1,7,8,16</sup> The Deyo adaptation of the Charlson comorbidity index and the individual comorbidities were used to measure overall comorbidity burden.<sup>17</sup> Additionally, obesity (ICD-9 codes: 278.0, 278.00, 278.01, 649.1, V85.3, V85.4, V85.54, 793.91) was considered.

Outcome variables were also determined based on previous studies.<sup>1,7</sup> Outcomes for the whole study cohort included wound complications (including infection, dehiscence, seroma, and hematoma), combined medical complications (including myocardial infarction, shock, deep vein thrombosis, pneumonia, pulmonary embolism, stroke, other central nervous system complication including hemorrhage, other unspecified complications not otherwise classified), utilization of spine imaging (including magnetic resonance imaging, computed tomography), admission to an intensive care unit (ICU), 30-day readmission, length of hospital stay (LOHS) in days, and cost of hospitalization (COH) in U.S. dollars.

Next to these general outcome variables, we also studied outcome variables for specific subgroups: (1) revisions in cervical spinal fusions; (2) revisions in thoracic fusions; (3) revisions in lumbar fusions; (4) dysphagia/hoarseness, emergency tracheostomy, and intubation in anterior cervical fusions. Similar to previous studies,<sup>1,7</sup> revision surgery was defined using ICD-9 codes 81.32 to 81.39, indicating refusion of the spine, without mention of the indication for revision

(e.g., pseudarthrosis). Complications were captured if they occurred during the index admission. We looked at refusion performed at either the index admission (the primary fusion procedure) or any subsequent admission within the study's time interval of 2006 to 2012. The admissions were linked by using a unique deidentified medical record key provided by Premier.

### Measures of Off-Label Use of BMP-2

Based on the FDA "Summary of Safety and Effectiveness Data" for BMP-2, we extracted the following variables depicting off-label use of BMP-2: (1) spinal fusions other than the lumbar anterior segment/approach, (2) diagnosis other than degenerative disk disease, (3)  $\geq 4$  levels of fusion, (4) patients  $< 18$  years, (5) malignant spine tumor, (6) pregnant women, (7) spinal abscess. The first four indicators were already included as study variables. Using ICD-9 coding, we were able to determine the presence of a malignant spine tumor (ICD-9 codes 192.2, 192.3, 198.3, 198.4, 237.5, 237.6), and patients likely to be pregnant at the time of surgery (ICD-9 coding from a previous study<sup>18</sup>). However, due to nonspecific ICD-9 codes, we were not able to reliably determine the presence of a spinal abscess. An additional measure of off-label BMP-2 use is its use with an interbody cage other than the LT cage (Medtronic, Memphis, Tennessee, United States). However, because we were not able to specify the device used, we could not use this measure of off-label use.

Next to these so-called "absolute" off-label measures for BMP-2 use, we also studied variables related to the "Warnings and Precautions" section in the FDA documentation; the relative contraindications include: (1) women of childbearing age (15 to 45 years), (2) patients age  $\geq 65$  years, (3) patients with any malignancy, (4) patients with renal compromise, and (5) patients with liver compromise. These indicators were already included as study variables.

### Statistical Analysis

#### Univariable Analysis

BMP-2 use was tabulated by covariates in the periods 2006 to 2008 and 2009 to 2012. The main reason for this dichotomization is the 2008 FDA safety warning.

Chi-square tests were used to compare BMP-2 use in 2006 to 2008 versus 2009 to 2012. Next, comorbidity burden and outcomes were tabulated between patients undergoing their procedures with use of BMP-2 versus those undergoing procedures without BMP-2 use in the two periods. Median and interquartile range were reported for LOHS and COH. Due to their skewed distribution, the significance between groups was measured using the Mann-Whitney rank sum test. To measure if a trend existed from 2006 to 2012 in the use of BMP-2, the Cochran Armitage trend test was performed separately for each procedure type.

#### Multivariable Analysis

Multilevel multivariable logistic regressions were performed to measure the effect of BMP-2 and the binary outcomes. The full study cohort was used for modeling the association

between BMP-2 use and wound complications, combined medical complications, utilization of spine imaging, admission to an ICU, and 30-day readmission. A selection of the study cohort was used for modeling the association between BMP-2 use and revisions (each separately for cervical, thoracic, and lumbar spinal fusions) or dysphagia/hoarseness and emergency tracheostomy and intubation (in anterior cervical fusions). We used multilevel modeling as the data has a hierarchical structure (i.e., patients within hospitals). A random intercept was included to account for the variation between hospitals. The multilevel modeling allows us to take into account the variation between hospitals, in addition to prespecified variables such as hospital teaching status, size, or rural/urban location.

To stratify by period, every model included an interaction between BMP-2 and period (either 2006 to 2008 or 2009 to 2012). All models were adjusted by all demographic, procedure, and health care–related covariates that were associated with BMP-2 use in the univariable analyses ( $p < 0.15$ ), which were age, gender, race, smoking status, insurance status, hospital location, hospital bed size, hospital teaching status, procedure diagnosis, segment/approach of fusion, number of vertebral levels fused, use of an interbody device, Deyo-Charlson Index, and obesity. Adjusted odds ratios (ORs), 95% confidence intervals (CIs), and  $p$  values are reported. CIs and  $p$  values were adjusted for multiple comparisons using the Bonferroni correction due to the interaction. Analyses were performed in SAS v9.3 statistical software (SAS Institute, Cary, North Carolina, United States).

### Results

The study sample consisted of 340,393 cases of elective primary cervical, thoracic, and lumbar spinal fusions from 418 hospitals.

#### Univariable Results

Overall, rates of BMP-2 use decreased from 18.7% of spinal fusions in 2006 to 11.5% in 2012 (►Table 1). The significant decrease in rates of BMP-2 use was seen for almost all patient subgroups. The mean Deyo-Charlson comorbidity index was only modestly higher for patients undergoing their spinal fusions with BMP-2 versus those having procedures without BMP-2 in both periods: 0.56 versus 0.60 and 0.67 versus 0.69 in BMP-2 versus no BMP-2 use in 2006 to 2008 and 2009 to 2012, respectively (both  $p < 0.001$ ; ►Table 2). Both ►Table 1 and ►Table 2 include variables depicting rates of off-label use of BMP-2. In almost all the groups, off-label rates decreased, most notably for the patients  $< 18$  years (9.4 versus 4.5%), cervical anterior fusions (5.1 versus 2.0%), and cervical posterior fusions (15.3 versus 8.5%; all  $p < 0.01$  comparing 2006 to 2008 with 2009 to 2012).

In 2006 to 2008, all the outcome variables were significantly higher in patients undergoing their procedures with use of BMP-2, except for 30-day readmission and the need for an emergency tracheostomy (►Table 3). Most differences remained in the 2009 to 2012 period; however, the comparison was significantly reversed for dysphagia/hoarseness for

**Table 1** BMP-2 use by covariates and time periods<sup>a</sup>

	2006–2008 (N = 132,789)		2009–2012 (N = 207,604)	
	No BMP-2	BMP-2	No BMP-2	BMP-2
	n (%)	n (%)	n (%)	n (%)
Overall	107,115 (80.7)	25,674 (19.3)	175,780 (84.7)	31,824 (15.3)
Patient demographics				
Age (y)				
<18 <sup>b</sup>	1,648 (90.6)	171 (9.4)	2,489 (95.5)	116 (4.5)
18–64	82,573 (81.1)	19,232 (18.9)	128,547 (85.1)	22,563 (14.9)
≥65 <sup>c</sup>	22,894 (78.5)	6,271 (21.5)	44,744 (83.0)	9,145 (17.0)
Gender				
Female	42,932 (79.7)	10,939 (20.3)	74,871 (83.9)	14,384 (16.1)
Of childbearing age (15–45 y) <sup>c</sup>	15,336 (81.7)	3,439 (18.3)	20,748 (85.4)	3,537 (14.6)
Pregnant <sup>b</sup>	10 (100)	0 (0.0)	11 (84.6)	2 (15.4)
Male	48,837 (81.2)	11,296 (18.8)	80,150 (85.2)	13,901 (14.8)
Race				
White	80,066 (81.1)	18,604 (18.9)	135,590 (84.6)	24,753 (15.4)
Black	8,270 (82.3)	1,776 (17.7)	14,892 (86.2)	2,382 (13.8)
Hispanic	2,548 (79.7)	647 (20.3)	2,346 (83.7)	456 (16.3)
Other	16,231 (77.7)	4,647 (22.3)	22,952 (84.4)	4,233 (15.6)
Smoker	30,446 (80.0)	7,597 (20.0)	51,241 (83.9)	9,814 (16.1)
Health care–related				
Insurance type				
Commercial	58,587 (81.6)	13,217 (18.4)	85,593 (85.5)	14,467 (14.5)
Medicaid	4,998 (82.2)	1,081 (17.8)	10,240 (86.6)	1,585 (13.4)
Medicare	29,587 (79.0)	7,878 (21.0)	59,135 (83.6)	11,608 (16.4)
Uninsured	1,290 (84.6)	234 (15.4)	2,005 (88.8)	253 (11.2)
Other	12,653 (79.5)	3,264 (20.5)	18,807 (82.8)	3,911 (17.2)
Hospital location				
Rural	10,092 (85.1)	1,768 (14.9)	16,497 (82.6)	3,477 (17.4)
Urban	97,023 (80.2)	23,906 (19.8)	159,283 (84.9)	28,347 (15.1)
Hospital size				
<300 beds	23,895 (83.1)	4,850 (16.9)	49,388 (84.5)	9,072 (15.5)
300–499 beds	43,071 (81.8)	9,582 (18.2)	68,058 (85.2)	11,852 (14.8)
≥500 beds	40,149 (78.1)	11,242 (21.9)	58,334 (84.3)	10,900 (15.7)
Hospital teaching status				
Nonteaching	63,086 (80.9)	14,894 (19.1)	104,682 (85.4)	17,921 (14.6)
Teaching	44,029 (80.3)	10,780 (19.7)	71,098 (83.6)	13,903 (16.4)
Mean annual no. of spinal fusions per hospital <sup>d</sup>	354 [233]	375 [238]	356 [254]	355 [230]
Procedure-related				
Diagnosis				
Degenerative disease	72,101 (84.2)	13,482 (15.8)	112,272 (88.6)	14,448 (11.4)
Spondylolisthesis <sup>b</sup>	2,114 (70.2)	896 (29.8)	3,247 (73.8)	1,155 (26.2)
Deformity <sup>b</sup>	2,067 (85.0)	366 (15.0)	3,219 (87.5)	459 (12.5)
Spinal stenosis <sup>b</sup>	5,072 (69.5)	2,225 (30.5)	9,835 (76.0)	3,104 (24.0)

(Continued)

**Table 1** (Continued)

	2006–2008 (N = 132,789)		2009–2012 (N = 207,604)	
	No BMP-2	BMP-2	No BMP-2	BMP-2
	n (%)	n (%)	n (%)	n (%)
Other diagnoses <sup>b</sup>	6,839 (80.7)	1,636 (19.3)	12,025 (87.7)	1,683 (12.3)
Multiple diagnoses	18,922 (72.8)	7,069 (27.2)	35,182 (76.2)	10,975 (23.8)
Segment/approach of fusion				
Cervical anterior <sup>b</sup>	55,487 (94.9)	3,003 (5.1)	88,245 (98.0)	1,818 (2.0)
Cervical posterior <sup>b</sup>	3,749 (84.7)	675 (15.3)	6,812 (91.5)	629 (8.5)
Thoracic anterior <sup>b</sup>	374 (74.4)	129 (25.6)	506 (81.9)	112 (18.1)
Thoracic posterior <sup>b</sup>	3,048 (82.8)	632 (17.2)	4,979 (82.2)	1,080 (17.8)
Lumbar anterior	3,451 (50.6)	3,363 (49.4)	6,979 (58.9)	4,868 (41.1)
Lumbar posterior and lateral transverse process <sup>b</sup>	40,870 (60.7)	17,767 (30.3)	68,003 (74.5)	23,236 (25.5)
Other <sup>b</sup>	136 (56.4)	105 (43.6)	256 (76.0)	81 (24.0)
Number of vertebral levels fused				
2–3	91,349 (80.7)	21,875 (19.3)	148,628 (84.5)	27,237 (15.5)
≥4 <sup>c</sup>	14,193 (80.7)	3,388 (19.3)	25,065 (85.2)	4,365 (14.8)
Unspecified	1,573 (79.3)	411 (20.7)	2,087 (90.4)	222 (9.6)
Use of interbody device				
Yes	47,140 (75.2)	15,530 (24.8)	93,949 (82.0)	20,560 (18.0)
No	59,975 (85.5)	10,144 (14.5)	81,831 (87.9)	11,264 (12.1)
Year of procedure				
2006	35,021 (81.3)	8,067 (18.7)	–	–
2007	35,834 (80.3)	8,805 (19.7)	–	–
2008	36,260 (80.5)	8,802 (19.5)	–	–
2009	–	–	42,728 (82.1)	9,325 (17.9)
2010	–	–	46,613 (82.9)	9,609 (17.1)
2011	–	–	48,015 (85.9)	7,874 (14.1)
2012	–	–	38,424 (88.5)	5,016 (11.5)

Abbreviation: BMP-2, bone morphogenetic protein-2.

<sup>a</sup>p Value between 0.01 and <0.0001 for comparison between rate of use in 2006 to 2008 versus 2009 to 2012 for all subgroupings except for (female) pregnant patients, chi-square test.

<sup>b</sup>Off-label BMP-2 use.

<sup>c</sup>Relative contraindications for BMP-2 use.

<sup>d</sup>Mean [standard deviation] instead of n (%); p < 0.0001 for comparison between 2006 to 2008 versus 2009 to 2012.

which those without BMP-2 had a risk of 3.0 versus 2.3% for those undergoing their procedure with use of BMP-2 ( $p < 0.001$ ). In particular, BMP-2 use was still significantly associated with increased LOHS (median 2 versus 3 days for both periods) and COH (median \$15,455 versus \$27,881 in 2006 to 2008 and \$17,007 versus \$30,331 in 2009 to 2012).

### Multivariable Results

When controlling for relevant covariates, the use of BMP-2 (in 2006 to 2008 and 2009 to 2012) was significantly associated with an increased risk of wound-related complications (OR 1.22 and 1.23), increased utilization of spine imaging (OR 1.23 and 1.11), increased risk for ICU admission (OR 1.40 and 1.36), increased risk of cervical revisions (OR 2.43 and 1.67), and increased risk for lumbar revisions (OR 1.85 and

1.55); **Table 4**. However, the 31% (OR 1.31) increased risk for dysphagia/hoarseness in 2006 to 2008 was not seen in 2009 to 2012. For most outcomes, there was a tendency of decreased ORs in 2009 to 2012 compared with 2006 to 2008, but with overlapping CIs. For all models, the c-statistic varied from 0.72 to 0.94, indicating good model discrimination.

### Discussion

Utilizing a specific and conservative indicator of BMP-2 use in a nationwide claims-based data set, largely untapped in spine research, our study shows a steady decrease in rates of overall and off-label use of BMP-2 in spinal fusions in the United States. Most notably, the decrease was observed in cervical spinal fusions. We further demonstrated a tendency of



**Table 2** Comorbidities by BMP-2 use in 2006 to 2009 and 2009 to 2012

	2006–2008			2009–2012		
	No BMP-2	BMP-2		No BMP-2	BMP-2	
	n (%)	n (%)	p Value <sup>a</sup>	n (%)	n (%)	p Value <sup>a</sup>
Mean Deyo-Charlson comorbidity index <sup>b</sup>	0.56 [0.95]	0.60 [0.97]	<0.001	0.67 [1.04]	0.69 [1.06]	<0.001
Deyo-Charlson comorbidity grouping						
Myocardial infarction	3,305 (3.1)	846 (3.3)	0.083	6,104 (3.5)	1,209 (3.8)	0.004
Congestive heart failure	1,807 (1.69)	502 (1.96)	0.003	3,339 (1.90)	668 (2.10)	0.017
Peripheral vascular disease	1,789 (1.7)	596 (2.3)	<0.001	3,632 (2.1)	827 (2.6)	<0.001
Cerebrovascular disease	261 (0.2)	60 (0.2)	0.770	484 (0.3)	91 (0.3)	0.740
Dementia	39 (0.04)	17 (0.1)	0.037	86 (0.05)	18 (0.1)	0.575
Chronic pulmonary disease	16,674 (15.6)	4,242 (16.5)	<0.001	30,057 (17.1)	5,691 (17.9)	0.001
Rheumatologic disease	2,505 (2.3)	762 (3)	<0.001	5,122 (2.9)	1,121 (3.5)	<0.001
Mild liver disease <sup>c</sup>	230 (0.2)	55 (0.2)	0.988	437 (0.2)	89 (0.3)	0.311
Moderate/severe liver disease <sup>c</sup>	48 (0.05)	12 (0.05)	0.896	143 (0.1)	20 (0.1)	0.278
Diabetes	15,126 (14.1)	3,781 (14.7)	0.013	29,279 (16.7)	5,380 (16.9)	0.273
Diabetes with chronic complications	1,236 (1.2)	340 (1.3)	0.024	2,641 (1.5)	508 (1.6)	0.208
Hemiplegia or paraplegia	214 (0.20)	65 (0.25)	0.093	495 (0.28)	86 (0.27)	0.724
Renal disease <sup>c</sup>	65 (0.1)	12 (0.05)	0.405	145 (0.1)	25 (0.1)	0.821
Any malignancy <sup>c</sup>	3,443 (3.2)	846 (3.3)	0.510	9,156 (5.2)	1,462 (4.6)	<0.001
Malignant spine tumor <sup>d</sup>	51 (0.05)	15 (0.1)	0.485	119 (0.1)	9 (0.03)	0.009
Metastatic solid tumor	423 (0.40)	76 (0.30)	0.020	776 (0.44)	64 (0.20)	<0.001
AIDS	36 (0.03)	12 (0.05)	0.320	94 (0.05)	13 (0.04)	0.361
Obesity	11,524 (10.8)	3,570 (13.9)	<0.001	26,320 (15.0)	5,382 (16.9)	<0.001

Abbreviation: AIDS, acquired immune deficiency syndrome; BMP-2, bone morphogenetic protein-2.

<sup>a</sup>Chi-square test for univariable association between Deyo-Charlson comorbidity grouping variables and BMP-2 use.

<sup>b</sup>Mean [standard deviation] instead of n (%).

<sup>d</sup>Off-label BMP-2 use.

<sup>c</sup>Relative contraindications for BMP-2 use.

decreased odds for adverse outcomes after BMP-2 use in 2009 to 2012 compared with 2006 to 2008. However, BMP-2 remains associated with increased resource utilization and ORs for adverse outcomes.

The decrease in BMP-2 use may reflect the mounting evidence on its association with increased risk of complications, as noted in the high-profile study by Cahill et al using data from over 300,000 patients.<sup>1</sup> It showed BMP-2 use in anterior cervical fusions to be associated with a >60% risk in wound-related complications and dysphagia/hoarseness (OR 1.67; 95% CI 1.10 to 2.53 and OR 1.63; 95% CI 1.30 to 2.05, respectively). We also found significantly increased adjusted odds for wound-related complications after BMP-2 use versus no BMP-2 use in the overall group of spinal fusions (► **Table 4**).

Another important event that may have played a major role in the decrease of BMP-2 use is the official warning issued in July of 2008 in which the FDA warned against BMP-2 use in treatment of cervical spine conditions and recommended alternative treatments for these patients or enrollment in approved clinical studies.<sup>10</sup> This warning followed from at least 38 events of

complications reported after BMP-2 use in cervical spine fusion in the previous 4 years. Similar to cases reported in the literature, these complications were associated with swelling of neck and throat tissue, which resulted in compression of the airway and/or neurologic structures in the neck.<sup>19,20</sup>

Next to these two significant events in the period since BMP-2 approval, there has also been an active discussion in the literature on authors' conflicts of interest in studies on the safety of BMP-2 use.<sup>5,11,21,22</sup> In 2011, Carragee et al compared industry trial results with those submitted to the FDA.<sup>11</sup> They showed that multiple industry-sponsored clinical trials studying BMP-2 use in spinal surgery underreported the rate of adverse events, even stating that the actual risk of adverse events for patients receiving BMP-2 was 10 to 50 times higher than reported in the industry-sponsored studies. Independent from this critical review and with better access to original study data, Rodgers et al came to similar conclusions.<sup>21</sup>

These developments highlighted the discussion on the safety of BMP-2 and substantially focused on adverse events after cervical spinal fusions, which in turn might have led to

**Table 3** Outcomes by BMP-2 use in 2006 to 2009 and 2009 to 2012

	2006–2008			2009–2012		
	No BMP-2	BMP-2		No BMP-2	BMP-2	
	<i>n</i> (%)	<i>n</i> (%)	<i>p</i> Value <sup>a</sup>	<i>n</i> (%)	<i>n</i> (%)	<i>p</i> Value <sup>a</sup>
All procedures						
Wound-related complications	1,089 (1)	394 (1.5)	<0.001	1,432 (0.8)	401 (1.3)	<0.001
Combined medical complications	2,679 (2.5)	1,013 (3.9)	<0.001	2,023 (1.2)	583 (1.8)	<0.001
Spine imaging	6,143 (5.7)	3,310 (12.9)	<0.001	26,751 (15.2)	7,021 (22.1)	<0.001
Admission to intensive care unit	15,654 (14.6)	4,346 (16.9)	<0.001	29,031 (16.5)	6,601 (20.7)	<0.001
30-d readmission	31 (0.03)	10 (0.04)	0.412	53 (0.03)	13 (0.04)	0.325
Length of hospital stay (d) <sup>b</sup>	2 [1–4]	3 [2–5]	<0.001	2 [1–3]	3 [2–4]	<0.001
Cost of hospitalization (\$) <sup>b</sup>	15,455 [10,334–23,895]	27,881 [20,957–37,362]	<0.001	17,007 [11,413–26,185]	30,331 [22,494–40,332]	<0.001
Selected procedures						
Cervical						
Revision	351 (0.6)	72 (1.9)	<0.001	744 (0.8)	46 (1.8)	<0.001
Thoracic						
Revision	51 (1.5)	23 (3)	0.004	99 (1.8)	52 (4.4)	<0.001
Lumbar						
Revision	487 (1.1)	364 (1.7)	<0.001	1,264 (1.7)	598 (2.1)	<0.001
Anterior cervical						
Dysphagia/hoarseness	1,323 (2.4)	160 (5.3)	<0.001	2,647 (3.0)	62 (2.3)	<0.001
Emergency tracheostomy	110 (0.2)	10 (0.3)	0.112	212 (0.2)	1 (0.1)	0.108
Emergency intubation	200 (0.4)	30 (1)	<0.001	411 (0.5)	12 (0.7)	0.230

Abbreviation: BMP-2, bone morphogenetic protein-2.

<sup>a</sup>Chi-square test for univariable association between outcome variables and BMP-2 use for categorical variables; Mann-Whitney Rank Sum for LOHS and COH.

<sup>b</sup>Median [interquartile range] instead of *n* (%).

increased awareness possibly leading to the observed changes in utilization of BMP-2 in terms of administration. Moreover, this change could have also affected other factors such as dosing (which we unfortunately could not study in our data set), which may reflect the observed changes in the tendency toward overall lower ORs for adverse outcomes in 2009 to 2012 versus the 2006 to 2008 period. We expect these changes have continued after 2012, especially because two highly anticipated meta-analyses were published in 2013 that resulted from the Yale University Open Data Access (YODA) Project.<sup>23–25</sup> Interestingly, YODA's key collaborator was Medtronic, BMP-2's manufacturer and main sponsor of early trials. The main aim was to perform a rigorous systematic review and evidence synthesis to determine safety and effectiveness of BMP-2. Medtronic provided all relevant patient-level clinical trial data to be included in this endeavor. Although the first meta-analysis concluded BMP-2 increased fusion rates, modestly reduced pain, and increased early

postsurgical pain, the second meta-analysis stated BMP-2 had no proven clinical advantage over bone graft and may be associated with important damage or injury. Thus, the challenge of identifying patients who are expected to benefit the most from the use of BMP-2 remains. Future studies that address this question should result in clearly defined patient selection criteria focusing on patients in whom the benefits of using BMP-2 exceed the potential harms.

Despite the overall decrease in rates of BMP-2 use, we still observed substantial off-label use of BMP-2. Most notably, 85 to 87% of procedures that used BMP-2 were performed in nonanterior lumbar fusions. This percentage is similar to that reported previously in a large-scale study using data from 2002 to 2007.<sup>8</sup> Although infrequent, of particular concern is the use of BMP-2 in patients with malignancies, because there is ongoing debate in the literature on the risk of novel neoplasms after the use of BMP-2.<sup>4,26</sup> Also, further research is needed on effectiveness and safety of BMP use (and its

**Table 4** Results from the multilevel logistic regression models for the binary outcomes in 2006 to 2008 and 2009 to 2012

	2006–2008	2009–2012	C statistic
	BMP-2 versus no BMP-2	BMP-2 versus no BMP-2	
	OR <sup>a</sup> (CI)	OR <sup>a</sup> (CI)	
All procedures			
Wound-related complications	1.22 (1.03–1.45) <sup>b</sup>	1.23 (1.04–1.46) <sup>c</sup>	0.72
Combined medical complications	1.13 (0.93–1.37)	0.94 (0.67–1.31)	0.78
Spine imaging	1.23 (1.13–1.34) <sup>c</sup>	1.11 (1.03–1.19) <sup>c</sup>	0.94
Admission to intensive care unit	1.40 (1.31–1.49) <sup>c</sup>	1.36 (1.28–1.44) <sup>c</sup>	0.88
30-d readmission <sup>d</sup>	–	–	–
Selected procedures			
Cervical			
Revision <sup>d</sup>	2.43 (1.66–3.54) <sup>c</sup>	1.67 (1.01–2.78) <sup>b</sup>	0.88
Thoracic			
Revision	1.49 (0.72–3.09)	1.59 (0.93–2.71)	0.82
Lumbar			
Revision	1.85 (1.50–2.28) <sup>c</sup>	1.55 (1.32–1.82) <sup>c</sup>	0.81
Anterior cervical			
Dysphagia/hoarseness	1.31 (1.02–1.67) <sup>b</sup>	0.94 (0.65–1.37)	0.88
Emergency tracheostomy	1.36 (0.52–3.56)	0.32 (0.21–4.77)	0.90
Emergency intubation	1.22 (0.70–2.11)	1.19 (0.52–2.71)	0.84

Abbreviations: BMP-2, bone morphogenetic protein-2; CI, confidence interval; OR, odds ratio.

<sup>a</sup>Adjusted for: age, gender, race, smoking status, insurance status, hospital location, hospital bed size, hospital teaching status, procedure diagnosis, segment/approach of fusion, number of vertebral levels fused, use of an interbody device, Deyo-Charlson Index, and obesity.

<sup>b</sup>p Value between 0.05 and 0.01.

<sup>c</sup>p Value between 0.01 and <0.0001.

<sup>d</sup>Too few events for 30-day readmissions; convergence issues for multilevel model with revision cervical spinal fusions as an outcome, effect estimates are derived from conventional multivariable logistic regression.

biological characteristics) in patients with relative contraindications such as old age or renal impairment,<sup>27,28</sup> as studies have put forward some concerns.

Similar to other studies, we too observed an increased use in resource utilization and hospital charges after spinal fusions with BMP-2 use versus no BMP-2 use.<sup>1,15</sup> A recently published study using 2002 to 2008 nationwide data for patients undergoing lumbar spinal fusions adjusted for other factors and showed that the use of BMP-2 was associated with a more than \$13,000 increase in COH.<sup>15</sup> Our study showed similar differences in costs in the univariable comparisons (► **Table 3**). Although the costs of BMP-2 itself have to be taken into account, along with likely confounding by procedure complexity and patient factors, this increase in overall costs might be related to indirect costs following from BMP-2 use (e.g., concerning its association with complications). However, the issue of cost-effectiveness is debated as some older and smaller studies resulted in conflicting conclusions.<sup>29,30</sup> Furthermore, the data on true cost of care, which would include the need for revision surgery in the setting of pseudarthrosis, is elusive on a large scale.

The main strengths of our analysis include the usage of recent data, the ability to show time trends, and the adjustment for hospital variation in the multilevel models. In

addition, unlike all previous large-scale studies on BMP-2 associations with adverse outcomes, we were the first to specifically identify BMP-2 use from billing data as opposed to studies using ICD-9 codes, which could comprise both BMP-2 and BMP-7.<sup>1,7,9,15</sup> This approach might have affected the results in these studies as some evidence suggests differential tissue reactions for both types of BMP.<sup>31</sup> Moreover, we observed that of the overall BMP use identified by ICD-9 coding, 55% was BMP-2 defined from billing definitions. This finding indicated ours to be a more conservative estimate of BMP-2 use compared with ICD-9 coding. Interestingly, in a previous study looking at 57 Medicare patients identified as having received BMP-2 by ICD-9 coding, chart review revealed a positive predictive value of 100% (all cases were confirmed to have received BMP-2).<sup>32</sup> We can only speculate what causes this discrepancy; however, as the number of studies on this topic is increasing, a study geared toward additionally validating ICD-9 coding or billing data for BMP-2 use (preferably multisite and all-payer) would greatly further this discussion. Moreover, the creation of a multisite registry geared toward the issue of BMP-2 safety might be of additional benefit. Another strength of the current study pertains to the fact that compared with previous studies,<sup>8</sup> we were able to broaden the discussion on off-



label use by using relative contraindications for BMP-2 use in addition to the more often mentioned absolute measures of off-label use.

Our study is burdened by several limitations. First, and most importantly, we utilized data from an administrative database, and detailed clinical information is missing. Examples are the instrumentation and use of grafting material, dosing, bone density, use of corticosteroids, the indications for revision surgery, and the apparent undercoding of dysphagia, which has been shown to have a positive predictive value of 94% but a negative predictive value of only 13%.<sup>23</sup> However, we do expect the undercoding to be similar across treatment groups. The use of administrative databases has several advantages, including (1) the rare availability of detailed and large-scale information on use of medications and medical devices allowing for analyses of this type and (2) the possibility to look into safety concerns regarding the incidence of rare complications. Moreover, as opposed to the highly selective patient population from clinical trials, administrative databases provide information from a wide range of hospitals and patients, thus representing actual practice settings. Second, although we adjusted for many important covariates in our models while also accounting for correlation of patients within hospitals, residual confounding remains, which may partly be caused by the lack of information on both the drivers for the use of BMP-2 and the indication for revision surgery. We expect the choice for BMP-2 to depend on numerous factors including severity of underlying disease, insurance approval, local hospital policies, or surgeon or patient preferences. These factors in turn may vary greatly between hospitals and even surgeons, which further stresses the need for studies looking into the current drivers of the choice to utilize BMP-2. The multilevel model may account in part for interhospital differences because it adjusts for unspecified effects on the hospital level (of which hospital-specific decision making regarding BMP-2 practice should be part). Next to lack of information on the drivers for BMP-2 utilization, residual confounding could also be caused by insufficient adjustment for surgical complexity, which we expect to be related to the decision for the use of BMP-2. However, we have tried to (partially) account for this observation by adjusting for diagnosis, use of an interbody device, and the levels of fusion. Although not a complete measure of case complexity, we expect a fair capture of the complexity with these variables.

Finally, we were not able to reliably gather information about the use of BMP-2 on previously mentioned long-term effects (e.g., neoplasms or retrograde ejaculation), which is also important in determining effectiveness and safety.<sup>6,33</sup> For example, we did not find an increased risk for retrograde ejaculation after anterior lumbar spinal fusion (data not shown), which was found to be the only significant complication associated with BMP-2 use in a recently published review.<sup>6</sup>

## Conclusions

In conclusion, using nationwide population-based data and a specific and conservative indicator of BMP-2 use, our study

shows a steady decrease in rates of overall and off-label use of BMP-2 in spinal fusions in the United States. Although associations with increased resource utilization and adverse outcomes remain, most outcomes have been positively affected, possibly due to increased awareness that might have led to changes in practice (e.g., in patient groups at high risk for complications such as patients undergoing anterior cervical fusions). This data might add to the current discussion on the adverse outcomes associated with BMP-2 use in spinal fusions as an alternative to iliac crest bone graft. A comprehensive evaluation of BMP-2 on clinical effectiveness, cost-effectiveness, and safety will require ongoing prospective studies or the creation of a multisite clinical registry with long-term follow-up of these patients.

## Disclosures

Jashvant Poeran: none

Mathias Opperer: none

Rehana Rasul: none

Madhu Mazumdar: none

Federico P. Girardi: Personal fees (Lanx, Nuvasive, Ortho-Development Corp., Depuy Spine); Stock ownership (Small Bone Innovations, Pioneer Surgical Technology, LifeSpan, Centinel Spine, Spinal Kinetics, Paradigm Spine); Consultant (Gerson Lehman Group, Spineart USA, OrthoDevelopment Corp., Depuy Spine); Speaking and/or teaching arrangements (PharmaWrite); Scientific advisory board (Scient'x USA, Spinal Kinetics, Centinel Spine, Spineart USA, HealthpointCapital, Paradigm Spine)

Alexander P. Hughes: Consultant (Nuvasive, Spineview); Grant (Nuvasive)

Stavros G. Memtsoudis: none

Vassilios Vougioukas: none

## Acknowledgments

Funding: Dr. Memtsoudis is funded by the Anna Maria and Stephen Kellen Career Development Award, New York, New York. Contributions of Dr. Mazumdar and Mrs. Rasul on this project were partly funded by the Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, New York.

## References

- 1 Cahill KS, Chi JH, Day A, Claus EB. Prevalence, complications, and hospital charges associated with use of bone-morphogenetic proteins in spinal fusion procedures. *JAMA* 2009;302(1): 58–66
- 2 Rajae SS, Bae HW, Kanim LE, Delamarter RB. Spinal fusion in the United States: analysis of trends from 1998 to 2008. *Spine (Phila Pa 1976)* 2012;37(1):67–76
- 3 U.S. Food and Drug Administration. InFuse™ Bone Graft/LT-CAGETM lumbar tapered fusion device-P000058, 2002. Available at: <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cftopic/pma/pma.cfm?num=P000058>. Accessed January 23, 2014
- 4 Devine JG, Dettori JR, France JC, Brodt E, McGuire RA. The use of rhBMP in spine surgery: is there a cancer risk? *Evid Based Spine Care J* 2012;3(2):35–41

- 5 Epstein NE. Complications due to the use of BMP/INFUSE in spine surgery: the evidence continues to mount. *Surg Neurol Int* 2013;4 (Suppl 5):S343–S352
- 6 Singh K, Ahmadinia K, Park DK, et al. Complications of spinal fusion with utilization of bone morphogenetic protein: a systematic review of the literature. *Spine (Phila Pa 1976)* 2014;39(1):91–101
- 7 Goode AP, Richardson WJ, Schectman RM, Carey TS. Complications, revision fusions, readmissions, and utilization over a 1-year period after bone morphogenetic protein use during primary cervical spine fusions. *Spine J* 2014;14(9):2051–2059
- 8 Ong KL, Villarraga ML, Lau E, Carreon LY, Kurtz SM, Glassman SD. Off-label use of bone morphogenetic proteins in the United States using administrative data. *Spine (Phila Pa 1976)* 2010;35(19):1794–1800
- 9 Fineberg SJ, Ahmadinia K, Oglesby M, Patel AA, Singh K. Hospital outcomes and complications of anterior and posterior cervical fusion with bone morphogenetic protein. *Spine (Phila Pa 1976)* 2013;38(15):1304–1309
- 10 U.S. Food and Drug Administration. FDA Public Health Notification: Life-threatening Complications Associated with Recombinant Human Bone Morphogenetic Protein in Cervical Spine Fusion, 2008. Available at: <http://www.fda.gov/medicaldevices/safety/alertsandnotices/publichealthnotifications/ucm062000.htm>. Accessed January 23, 2014
- 11 Carragee EJ, Hurwitz EL, Weiner BK. A critical review of recombinant human bone morphogenetic protein-2 trials in spinal surgery: emerging safety concerns and lessons learned. *Spine J* 2011;11(6):471–491
- 12 U.S. Food and Drug Administration. OP-1TM-H010002, 2001. Available at: <http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/DeviceApprovalsandClearances/Recently-ApprovedDevices/ucm085026.htm>. Accessed January 23, 2014
- 13 Premier, Inc. Premier Perspective Database. Available at: <https://www.premierinc.com/transforming-healthcare/healthcare-performance-improvement/premier-research-services/>. Accessed January 2, 2014
- 14 Rothberg MB, Pekow PS, Priya A, et al. Using highly detailed administrative data to predict pneumonia mortality. *PLoS ONE* 2014;9(1):e87382
- 15 Dagostino PR, Whitmore RG, Smith GA, Maltenfort MG, Ratliff JK. Impact of bone morphogenetic proteins on frequency of revision surgery, use of autograft bone, and total hospital charges in surgery for lumbar degenerative disease: review of the Nationwide Inpatient Sample from 2002 to 2008. *Spine J* 2014;14(1):20–30
- 16 Pumberger M, Chiu YL, Ma Y, Girardi FP, Vougioukas V, Memtsoudis SG. Perioperative mortality after lumbar spinal fusion surgery: an analysis of epidemiology and risk factors. *Eur Spine J* 2012;21(8):1633–1639
- 17 Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 1992;45(6):613–619
- 18 Jin A, Martin JD, Sarin C. Diabetes mellitus in the First Nations population of British Columbia, Canada. Part 2. Hospital morbidity. *Int J Circumpolar Health* 2002;61(3):254–259
- 19 Perri B, Cooper M, Laurusen C, Anand N. Adverse swelling associated with use of rh-BMP-2 in anterior cervical discectomy and fusion: a case study. *Spine J* 2007;7(2):235–239
- 20 Vaidya R, Carp J, Sethi A, Bartol S, Craig J, Les CM. Complications of anterior cervical discectomy and fusion using recombinant human bone morphogenetic protein-2. *Eur Spine J* 2007;16(8):1257–1265
- 21 Rodgers MA, Brown JV, Heirs MK, et al. Reporting of industry funded study outcome data: comparison of confidential and published data on the safety and effectiveness of rhBMP-2 for spinal fusion. *BMJ* 2013;346:f3981
- 22 Mirza SK. Folly of FDA-approval studies for bone morphogenetic protein. *Spine J* 2011;11(6):495–499
- 23 Krumholz HM, Ross JS. A model for dissemination and independent analysis of industry data. *JAMA* 2011;306(14):1593–1594
- 24 Simmonds MC, Brown JV, Heirs MK, et al. Safety and effectiveness of recombinant human bone morphogenetic protein-2 for spinal fusion: a meta-analysis of individual-participant data. *Ann Intern Med* 2013;158(12):877–889
- 25 Fu R, Selph S, McDonagh M, et al. Effectiveness and harms of recombinant human bone morphogenetic protein-2 in spine fusion: a systematic review and meta-analysis. *Ann Intern Med* 2013;158(12):890–902
- 26 Carragee EJ, Chu G, Rohatgi R, et al. Cancer risk after use of recombinant bone morphogenetic protein-2 for spinal arthrodesis. *J Bone Joint Surg Am* 2013;95(17):1537–1545
- 27 Lee KB, Taghavi CE, Hsu MS, et al. The efficacy of rhBMP-2 versus autograft for posterolateral lumbar spine fusion in elderly patients. *Eur Spine J* 2010;19(6):924–930
- 28 Latzman JM, Kong L, Liu C, Samadani U. Administration of human recombinant bone morphogenetic protein-2 for spine fusion may be associated with transient postoperative renal insufficiency. *Spine (Phila Pa 1976)* 2010;35(7):E231–E237
- 29 Garrison KR, Donell S, Ryder J, et al. Clinical effectiveness and cost-effectiveness of bone morphogenetic proteins in the non-healing of fractures and spinal fusion: a systematic review. *Health Technol Assess* 2007;11(30):1–150, iii–iv
- 30 Glassman SD, Carreon LY, Campbell MJ, et al. The perioperative cost of Infuse bone graft in posterolateral lumbar spine fusion. *Spine J* 2008;8(3):443–448
- 31 Lee KB, Taghavi CE, Murray SS, Song KJ, Keorochana G, Wang JC. BMP induced inflammation: a comparison of rhBMP-7 and rhBMP-2. *J Orthop Res* 2012;30(12):1985–1994
- 32 Mines D, Gu Y, Kou TD, Cooper GS. Recombinant human bone morphogenetic protein-2 and pancreatic cancer: a retrospective cohort study. *Pharmacoepidemiol Drug Saf* 2011;20(2):111–118
- 33 Even J, Eskander M, Kang J. Bone morphogenetic protein in spine surgery: current and future uses. *J Am Acad Orthop Surg* 2012;20(9):547–552