WASHINGTON STATE HEALTH CARE AUTHORITY

APPENDICES

Health Technology Assessment

Tuesday, February 14, 2012

Health Technology Assessment Program
676 Woodland Square Loop SE
P.O. Box 42712
Olympia, WA 98504-2712
http://www.hta.hca.wa.gov

On- and off-label uses of rhBMP-2 or rhBMP-7 for spinal fusion

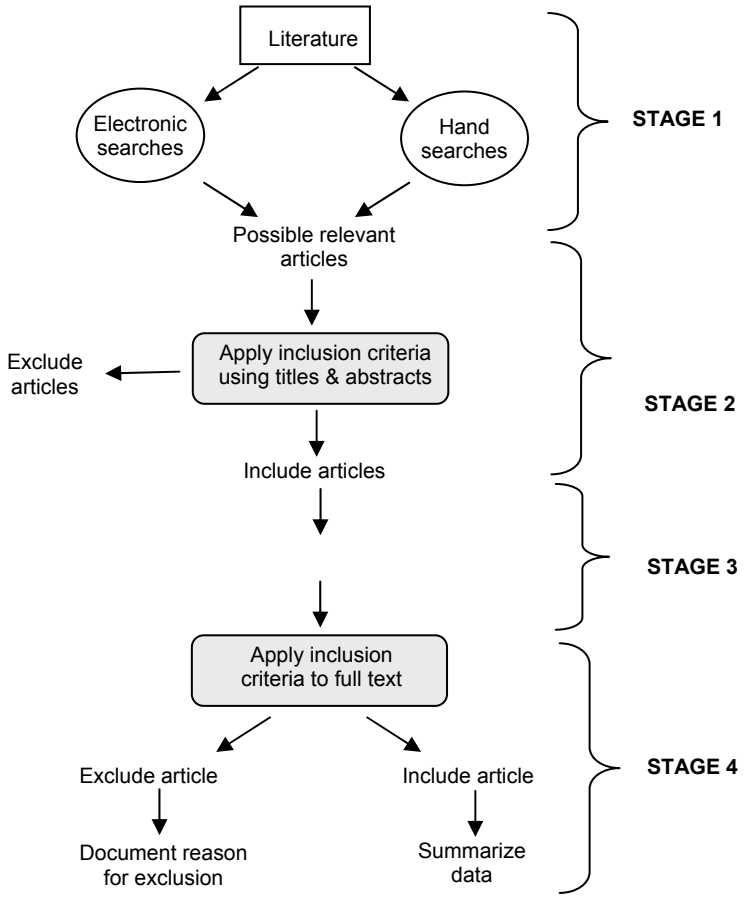
Provided by:

Spectrum Research, Inc.

APPENDICES

TABLE OF CONTENTS

Appendix A. ALGORITHM FOR ARTICLE SELECTION



Appendix B. SEARCH STRATEGIES

Key Question 1

Database: MEDLINE
Limit: English, only items with abstracts

	Limits: English, only items with abstracts	
	Search Terms	**No. of Articles**
#1	"Spinal Fusion"[MeSH] OR "spinal fusion" OR fusion*	178654
#2	(ODI OR "Oswestry Disability")	1488
#3	(valid* OR reliable OR reliability)	456595
#4	#1 AND #2 AND #3	45
#5	(NDI OR "Neck Disability Index")	1070
#6	#1 AND #3 and #5	16
#7	(VAS OR "Visual Analog Pain Scale" OR "Visual Analogue Pain Scale")	19705
#8	#1 AND #3 and #7	48
#9	(SF-36 OR "Short form 36")	11434
#10	#1 AND #3 and #9	42
#11	"minimal clinically important difference" OR "minimal important change"	263
#12	#1 AND #11	9

Total number of articles retrieved from search: 160
Total number of articles retrieved from handsearching for related references: 30
Total number of articles identified for review: 196

Key question 2-3: comparative studies

Note: the search for comparative studies was performed to identify only studies published after the search period used in the AHRQ HTA on BMP, as we accepted the search results from that HTA to identify comparative studies.

Search date: 9/14/2011

Limits: English, only items with abstracts, publication date 01/01/2010 (slight overlap with end of AHRQ search period) or later

	Search	Number of articles
#1	"Spinal Fusion"[MeSH] OR "spinal fusion" OR fusion*	16003
#2	"therapeutic use"[Subheading] OR "surgery"[Subheading] OR "injuries"[Subheading]	194341
#3	fracture* OR non-union* OR nonunion* OR fusion* OR allograft* OR autograft* OR arthrodes* OR malunion*	32876
#4	#1 OR #2 OR #3	216636
#5	"Bone Morphogenetic Proteins"[MeSH]	1375
#6	"bone morphogenetic" OR BMP OR BMP-2 OR BMP2 OR BMP-7 OR BMP7 OR rhBMP or rhBMP-2 OR rhBMP2 OR rhBMP-7 OR rhBMP7 OR rh-BMP or rh-BMP-2 OR rh-BMP2 OR rh-BMP-7 OR rh-BMP7 OR RHOP OR RHOP-1 OR op-1 OR op1	2794
#7	#5 OR #6	2796
#8	#4 AND #7	628
#10	#8 NOT (Animals[MeSH] OR "Models, Animal"[MeSH] OR "in vivo"[ti] OR "in vitro"[ti] NOT "Humans"[MeSH])	424
#12	#9 NOT (dental OR dentin OR odont* OR endodont* OR tooth OR teeth OR periodont* OR alveolar* OR cranio*[ti] OR calvaria*[ti] OR crania*[ti] OR jaw[ti] OR facial[ti] OR maxillofacia*[ti] OR maxilla-facia*[ti] OR mandib*[ti])	366
#14	#12 NOT ("Case Reports" [Publication Type])	346

22 additional studies included from AHRQ's HTA on BMP
Total number of articles evaluated for KQ2: 368

Key question 3: search for non-comparative studies
(see above for search for comparative studies)

Case series and case reports were identified using the following search:

Search date: 9/12/2011
Limits: English, only items with abstracts, publication date 01/01/1998

	Search	Number of articles
#1	"Spinal Fusion"[MeSH] OR "spinal fusion" OR fusion*	118211
#2	"therapeutic use"[Subheading] OR "surgery"[Subheading] OR "injuries"[Subheading]	1394350
#3	fracture* OR non-union* OR nonunion* OR fusion* OR allograft* OR autograft* OR arthrodes* OR malunion*	209354
#4	#1 OR #2 OR #3	1531322
#5	"Bone Morphogenetic Proteins"[MeSH]	8688
#6	"bone morphogenetic" OR BMP OR BMP-2 OR BMP2 OR BMP-7 OR BMP7 OR rhBMP or rhBMP-2 OR rhBMP2 OR rhBMP-7 OR rhBMP7 OR rh-BMP or rh-BMP-2 OR rh-BMP2 OR rh-BMP-7 OR rh-BMP7 OR RHOP OR RHOP-1 OR op-1 OR op1	13035
#7	#5 OR #6	13039
#8	#4 AND #7	3145
#9	#8 NOT (Animals[MeSH] OR "Models, Animal"[MeSH] OR "in vivo"[ti] OR "in vitro"[ti] NOT "Humans"[MeSH])	1907
#10	#9 NOT (dental OR dentin OR odont* OR endodont* OR tooth OR teeth OR periodont* OR alveolar* OR cranio*[ti] OR calvaria*[ti] OR crania*[ti] OR jaw[ti] OR facial[ti] OR maxillofacia*[ti] OR maxilla-facia*[ti] OR mandib*[ti])	1568
#11	"adverse events" OR "adverse event" OR "adverse effects"[subheading] OR antibody OR antibodies OR "allergic reaction" OR "allergic reactions" OR "Bone Morphogenetic Proteins/adverse effects"[MeSH] OR "Bone Transplantation*/adverse effects" OR "cancer" OR "cancers" OR "cerebrospinal fluid leak" OR "Cerebrospinal fluid leak"[Supplementary Concept] OR "Cervical Vertebrae/drug effects"[Mesh] OR complication* OR cardiac OR cardiovascular OR dehiscence OR death OR deaths OR Death[MeSH] OR "deep vein thrombosis" OR "Venous Thrombosis"[MESH] OR "durotomy" OR "durotomies" OR "dural tear" OR "dural tears" OR "delayed radiculopathy" OR displacement OR dysphagia OR "Deglutition Disorders"[MeSH] OR "ectopic bone formation" OR "ectopic ossification" OR "graft migration" OR "graft site morbidity" OR "graft site pain" OR hematoma* OR "hematoma"[MeSH] OR "heterotopic bone formation" OR "heterotopic ossification" OR Hypersensitivity[MeSH] OR infection OR infections OR "Infection"[MeSH] OR "Lumbar Vertebrae/drug effects"[Mesh] OR "Lumbosacral Region/drug effects"[Mesh] OR malposition* OR misposition* OR malignant OR malignancies OR "Neoplasms"[MeSH] OR neoplasm* OR osteolysis OR "Off-Label Use"[MAJR] OR "Osseointegration/drug effects*" OR	3124465

	"Spondylolisthesis/complications" OR "Postoperative Complications"[MAJR] OR "paresis" OR "Paresis"[MeSH] OR pseudarthrosis OR "Pseudarthrosis"[MeSH] OR resorption OR "retrograde ejaculation" OR reoperation* OR "Reoperation"[MeSH] OR revision* OR repair OR repairs* OR "Sacrum/drug effects"[Mesh] OR "Spinal Fusion/ adverse effects " OR "Safety"[Mesh] OR "safety" OR "Safety-Based Medical Device Withdrawals"[Mesh] OR subsidence OR swelling OR "surgical wound infection"[MeSH] OR sepsis OR "Sepsis"[MeSH] OR seroma* OR "seroma"[MeSH] OR "Surgical Wound Dehiscence"[MeSH] OR toxic OR toxicity OR "toxicity" [Subheading] OR tears* OR "urogenital"	
#12	#10 AND #11 AND ("Case Reports" [Publication Type] OR "case report" OR "case series" OR "series" OR consecutive OR "evaluation studies" OR "evaluation study" OR "retrospective evaluation" OR "retrospective studies" OR "retrospective studies"[MeSH] OR "present series" OR "retrospective review" OR "follow-up studies"[MeSH])	242
	Additional articles identified by handsearching bibliographies of included studies and by searching for related studies on Pubmed of included studies	1
	Total number of studies reviewed for inclusion by title/abstract:	**243***

* we verified that all relevant noncomparative studies in the AHRQ HTA on BMP were identified.

Key question 4:

For Key Question 4, evidence that the effects of treatment varied by sociological or demographic subgroups, we examined for inclusion the 44 comparative studies evaluated in Key Questions 2 and 3 (see above and methods section for details). Randomized controlled trials and non-randomized observational studies with concurrent controls evaluating surgical fusion versus non-operative management for chronic LBP were considered. The following criteria were used for inclusion in KQ4: RCTs that stratified the random assignment on one or more sociological or demographic subgroups, RCTs or non-randomized observational studies that included a subgroup analysis stratifying on one or more sociological or demographic subgroups, and RCTs or non-randomized observational studies that compared treatment among patients within specific sociological or demographic subgroups (e.g., older patients only) to compare with other comparative studies conducted among patients with the specific sociological or demographic subgroups (e.g., primarily younger patients). We excluded case series that provided subgroup analysis of sociological or demographic variables because this study design does not address the question of whether treatment differences vary according to differing sociological or demographic characteristic[1-4]. Articles were also excluded if they were pediatric studies (< 18 years of age), non-fusion surgeries, tumor surgery, revision surgery, treatment for osteomyelitis or trauma. Other exclusions included reviews, editorials, case reports, and non-English written studies, and studies without subgroup analyses.

Number of studies retrieved/evaluated at each step.

No. of studies from Pubmed search	n/a (we looked at FT for all comparative studies (RCTs, cohort) and database studies)
No. of studies identified from handsearching	n/a
Total number of studies identified for review (RCTs, cohort, and database studies with concurrent controls)	44
Number of studies retrieved for full text evaluation	44
Total number of studies excluded at full text evaluation	36
Total number of studies included in KQ4	8

Key question 5:

Search date: September, 2011
Limits**:** English, only items with abstracts, publication date starting 01/01/1998

	Search	Number of articles
#1	"Spinal Fusion"[MeSH] OR "spinal fusion" OR fusion AND "cost effectiveness"	226
#2	"Bone Morphogenetic Proteins"[MeSH] AND spinal fusion AND cost effectiveness	10

236 articles in total evaluated for inclusion

Parallel strategies were used to search the Cochrane Library and others listed below. Keyword searches were conducted in the other listed resources.

Electronic Database Searches
The following databases have been searched for relevant information (through August, 2011):
Agency for Healthcare Research and Quality (AHRQ)
Cumulative Index to Nursing and Allied Health (CINAHL)
Cochrane Database of Systematic Reviews
Cochrane Registry of Clinical Trials (CENTRAL)
Cochrane Review Methodology Database
Computer Retrieval of Information on Scientific Projects (CRISP)
Database of Reviews of Effectiveness (Cochrane Library)
EMBASE (1985 through August, 2010)
PubMed (1975 through August, 2010)
Informational Network of Agencies for Health Technology Assessment (INAHTA)
NHS Economic Evaluation Database
HSTAT (Health Services/Technology Assessment Text)
EconLIT

Additional Economics, Clinical Guideline and Gray Literature Databases
AHRQ- Healthcare Cost and Utilization Project
Canadian Agency for Drugs and Technologies in Health
Centers for Medicare and Medicaid Services (CMS)
Food and Drug Administration (FDA)
Google
Institute for Clinical Systems Improvement (ICSI)
National Guideline Clearinghouse

Appendix C. EXCLUDED ARTICLES

Exclude at full-text review

KQ1

Author	Year	Reason for exclusion
1. Blount	2002	Review article
2. Carragee	2010	No MCID values, just a minimal acceptable outcome
3. Carreon	2011	Algorithm for prediction of SF-6D from NDI
4. Carreon	2009	Algorithm for prediction of SF-6D from ODI
5. Cortes	2010	EQ-5D VAS, not pain VAS
6. Davidson	2002	Used a subset of SF-36
7. Donaldson	2011	Did not evaluate ODI, NDI, or SF-36
8. Helenius	2005	Evaluated SRS-30
9. Resnick	2005	Review article
10. Schwab	2008	Evaluated predictive models
11. Skolasky	2011	Predictive model for CSOQ
12. Skolasky	2007	Predictive model for CSOQ
13. Svensson	2009	Evaluated BIS

KQ2

Study	Reason for exclusion
BMP-2 off-label use (lumbar spine)	
1. Good 2010	Compares A/P to posterior only fusion; some patients in each treatment group received BMP

KQ3

Study	Reason for exclusion
BMP-2 on-label use (lumbar spine)	
1. Burkus 2003	No additional safety data reported (subset of Burkus 2002 RCT)
2. Kellman 2001	No safety data reported
BMP-2 off-label use (lumbar spine)	
3. Burkus 2002	Subset of Burkus 2005
4. Maeda 2009	All patients reported in Crawford 2010
5. Good 2010	Compares A/P to posterior only fusion; some patients in each treatment group received BMP
6. Rihn 2009 "use of"	Subset of patients reported in Rihn 2009 "complications"
7. Hamilton 2010	Subset of patients reported in Hamilton 2011

KQ4

Author	Year	Reason for exclusion
RCTs		
1. Baskin	2003	No subgroup analysis
2. Boden	2000	No subgroup analysis
3. Boden	2002	No subgroup analysis
4. Burkus	2002	No subgroup analysis
5. Burkus	2003	No subgroup analysis
6. Burkus	2005	No subgroup analysis
7. Dawson	2009	No subgroup analysis
8. Delawi	2010	No subgroup analysis
9. Dimar	2009	No subgroup analysis
10. Glassman	2008	No subgroup analysis
11. Haid	2004	No subgroup analysis
12. Hwang	2010	No subgroup analysis
13. Johnsson	2002	No subgroup analysis
14. Kanayama	2006	No subgroup analysis
15. Vaccaro	2004/2005/2008	No subgroup analysis
16. Vaccaro, Lawrence	2008	No subgroup analysis
Database Studies		
17. Mines	2011	No subgroup analysis
18. Cahill	2011	No subgroup analysis
Cohort Studies		
19. Burkus	2011	No subgroup analysis
20. Burkus, Sandhu	2066	No subgroup analysis
21. Buttermann	2008	No subgroup analysis
22. Crawford	2009	No subgroup analysis
23. Crawford	2010	No subgroup analysis
24. Howard	2011	No subgroup analysis
25. Joseph	2007	No subgroup analysis
26. Latzman	2010	No subgroup analysis
27. Lee	2010	No subgroup analysis
28. Maeda	2009	No subgroup analysis
29. Mummaneni	2004	No subgroup analysis
30. Pradhan	2006	No subgroup analysis
31. Singh	2006	No subgroup analysis
32. Smucker	2006	No subgroup analysis
33. Vaidya and Weir	2007	No subgroup analysis
34. Vaidya Carp	2007	No subgroup analysis
35. Xu	2011	No subgroup analysis
36. Yaremchuk	2011	No subgroup analysis

KQ5

Study	Reason for exclusion
1. Cahill 2009	Examines associations of rhBMP and hospital charges (regression analyses); does not include cost-effectiveness analysis models.
2. Polly 2003	Not a cost effectiveness study
3. Buttermann 2008	Examines outcomes of a cohort study, includes costs of treatments but does not include cost-effectiveness analysis models.
4. Glassman 2008	Intended to assess hospital costs, not to analyze clinical effectiveness or cost-effectiveness.
5. Ackerman 2002	Discusses methodology for cost analyses, not cost effectiveness study.
6. Cardoso 2009	Summarizes issues associated with determining whether this treatment is cost effective.

Appendix D. LEVEL AND STRENGTH OF EVIDENCE DETERMINATION

Methods for critical appraisal and level of evidence assessment

The method used for assessing the quality of evidence of individual studies as well as the overall quality of evidence incorporates aspects of rating scheme developed by the Oxford Centre for Evidence-based Medicine[5], precepts outlined by the Grades of Recommendation Assessment, Development and Evaluation (GRADE) Working Group[6] and recommendations made by the Agency for Healthcare Research and Quality (AHRQ)[7]. Taking into account features of methodological quality and important sources of bias combines epidemiologic principles with characteristics of study design.

Procedures for determining adherence to level of evidence (LoE) criteria
Each study was rated against pre-set criteria that resulted in an evidence rating (Level of Evidence I, II (IIa or IIb), III, or IV) and presented in a table. For therapeutic and prognostic articles, the criteria are listed in the Table below. All criteria met are marked. A "+" signifies that the criterion was present, a "–" indicates that the criterion was not present, and "+/–" indicates that the reviewers could not be determine whether the criterion was met.

After the Level of Evidence was judged, the study could be upgraded or downgraded using the following:

Upgrade: Large effect size, dose response
Downgrade: limitations in study execution, indirectness of evidence

Definition of the different levels of evidence for articles on therapy and prognosis

	Studies of Therapy		Studies of Prognosis	
Level	**Study design**	**Criteria**	**Study design**	**Criteria**
I	Good quality RCT	• Random sequence generation • Allocation concealment • Intent-to-treat analysis • Blind or independent assessment for important outcomes • Co-interventions applied equally • F/U rate of 80%+ • Adequate sample size	Good quality cohort	• Prospective design • Patients at similar point in the course of their disease or treatment • F/U rate of 80%+ • Patients followed long enough for outcomes to occur • Controlling for extraneous prognostic factors*
II	Moderate (IIa) or Poor (IIb) quality RCT	• Violation of one of the criteria for good quality RCT • Violation of two or more criteria for a good quality RCT	Moderate quality cohort	• Prospective design, with violation of one of the other criteria for good quality cohort study • Retrospective design, meeting all the rest of the criteria in level I
	Good quality cohort	• Blind or independent assessment in a prospective study, or use of reliable data* in a retrospective study • Co-interventions applied equally • F/U rate of 80%+ • Adequate sample size • Controlling for possible confounding†		
III	Moderate or poor quality cohort	• Violation of any of the criteria for good quality cohort	Poor quality cohort	• Prospective design with violation of 2 or more criteria for good quality cohort, or • Retrospective design with violation of 1 or more criteria for good quality cohort
	Case-control	• Any case-control design	Case-control	• Any case-control design
IV	Case series	• Any case series design	Case series	• Any case series design

*Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.

Determination of Overall Strength of Evidence

Following the assessment of the quality of each individual study included in the report, an overall "strength of evidence for the relevant question or topic is determined. Methods for determining the overall strength of evidence for diagnostic studies are variable across the literature and are most applicable to evaluation of therapeutic studies.

SRI's method incorporates the primary domains of quality (LoE), quantity of studies and consistency of results across studies as described by AHRQ[8].

SRI establishes a strength-of-evidence baseline using the following definitions to determine whether or not the body or evidence meets the criteria for each domain:

Domain	Definition/Criterion
Quality	• At least 80% of the studies are LoE I or II
Quantity	• There are at least three studies which are adequately powered to answer the study question
Consistency	• Study results would lead to a similar conclusion (similar values, in the same direction) in at least 70% of the studies (assumes at least three studies are available)

Based on the criteria described above, the possible scenarios that would be encountered are described below. Each scenario is ranked according to the impact that future research is likely to have on both the overall estimates of an effect and the confidence in the estimate. This ranking describes the overall "Strength of Evidence" (SoE) for the body of literature on a specific topic. The method and descriptions of overall strength are adapted for diagnostic studies from system described by the GRADE Working Group[6] for the development of clinical guidelines.

SoE	Description	Further Research Impact	Domain Criterion Met		
			Quality	Quantity	Consistency
1	**High**	Very unlikely to change confidence in effect estimate	+	+	+
2	**Moderate**	Likely to have an important impact on confidence in estimate and *may* change the estimate	+	-	+
			+	+	-
3	**Low**	Very likely to have an important impact on confidence in estimate and *likely* to change the estimate	+	-	-
			-	+	+
4	**Insufficient**	Any effect estimate is uncertain	-	+	-
			-	-	+
			-	-	-

Limitations or special strengths can modify the quality of the evidence from the baseline as follows:

Factors that can reduce the quality of the evidence 1 or 2 levels:

- Limitations in study design or execution
- Indirectness of evidence
- Imprecision

Factors that can increase the quality of the evidence: 1 or 2 levels:

- Large magnitude of effect
- Dose response gradient

Assessment of Economic Studies

Full formal economic analyses evaluate both costs and clinical outcomes of two or more alternative interventions. The four primary types are cost minimization analysis (CMA), cost-utility analysis (CUA), cost-effectiveness analysis (CEA), and cost-benefit analyses (CBA). Each employs different methodologies, potentially complicating critical appraisal, but some common criteria can be assessed across studies.

No standard, universally accepted method of critical appraisal of economic analyses is currently in use. A number of checklists [Canadian, BMJ, AMA] are available to facilitate critique of such studies. The Quality of Health Economic Studies (QHES) instrument developed by Ofman, et al[9]. QHES embodies the primary components relevant for critical appraisal of economic studies[9, 10]. It also incorporates a weighted scoring process and which was used as one factor to assess included economic studies. This tool has not yet undergone extensive evaluation for broader use but provides a valuable starting point for critique.

In addition to assessment of criteria in the QHES, other factors are important in critical appraisal of studies from an epidemiologic perspective to assist in evaluation of generalizability and potential sources of study bias.

Such factors include:

- Are the interventions applied to similar populations (eg, with respect to age, gender, medical conditions, etc)? To what extent are the populations for each intervention comparable and are differences considered or accounted for? To what extent are population characteristics consistent with "real world" applications of the comparators?
- Are the sample sizes adequate so as to provide a reasonable representation of individuals to whom the technology would be applied?
- What types of studies form the basis for the data used in the analyses? Data (eg, complication rates) from randomized controlled trials or well-conducted, methodologically rigorous cohort studies for data collection are generally of highest quality compared with case series or studies with historical cohorts.
- Were the interventions applied in a comparable manner (eg, similar protocols, follow-up procedures, evaluation of outcomes, etc)?
- How were the data and/or patients selected or sampled (eg, a random selection of claims for the intervention from a given year/source or all claims)? What specific inclusion/exclusion criteria or processes were used?
- Were the outcomes and consequences of the interventions being compared comparable for each? (eg, were all of the relevant consequences/complications for each intervention considered or do they primarily reflect those for one intervention?)

Assessment of the overall strength of evidence for formal economic analyses does not appear to be documented in the literature. For the purposes of this HTA, overall strength was determined by:

- Quality of the individual studies: Where the majority of quality indicators described in the QHES met and were the methods related to patient/claim selection, patient population considerations and other factors listed above consistent with a high quality design?
- Number of formal analyses (3 or more)
- Consistency of findings and conclusions from analyses across studies.

QHES Instrument[9] **Study** AHRQ HTA - Cost effectiveness analysis

Questions	Possible Points	Points Awarded
1. Was the study objective presented in a clear, specific, and measurable manner?	7	7
2. Were the perspective of the analysis (societal, third-party payer, etc.) and reasons for its selection stated?	4	4
3. Were variable estimates used in the analysis from the best available source (ie, randomized controlled trial - best, expert opinion - worst)?	8	6
4. If estimates came from a subgroup analysis, were the groups prespecified at the beginning of the study?	1	1
5. Was uncertainty handled by (1) statistical analysis to address random events, (2) sensitivity analysis to cover a range of assumptions?	9	9
6. Was incremental analysis performed between alternatives for resources and costs?	6	6
7. Was the methodology for data abstraction (including the value of health states and other benefits) stated?	5	5
8. Did the analytic horizon allow time for all relevant and important outcomes? Were benefits and costs that went beyond 1 year discounted (3% to 5%) and justification given for the discount rate?	7	7
9. Was the measurement of costs appropriate and the methodology for the estimation of quantities and unit costs clearly described?	8	8
10. Were the primary outcome measure(s) for the economic evaluation clearly stated and did they include the major short-term, long-term and negative outcomes included?	6	5
11. Were the health outcomes measures/scales valid and reliable? If previously tested valid and reliable measures were not available, was justification given for the measures/scales used?	7	7
12. Were the economic model (including structure), study methods and analysis, and the components of the numerator and denominator displayed in a clear, transparent manner?	8	8
13. Were the choice of economic model, main assumptions, and limitations of the study stated and justified?	7	7
14. Did the author(s) explicitly discuss direction and magnitude of potential biases?	6	4
15. Were the conclusions/recommendations of the study justified and based on the study results?	8	8
16. Was there a statement disclosing the source of funding for the study?	3	3
TOTAL POINTS	**100**	**95**

QHES Instrument[9] **Study** Carreon

Questions	Possible Points	Points Awarded
1. Was the study objective presented in a clear, specific, and measurable manner?	7	7
2. Were the perspective of the analysis (societal, third-party payer, etc.) and reasons for its selection stated?	4	2
3. Were variable estimates used in the analysis from the best available source (ie, randomized controlled trial - best, expert opinion - worst)?	8	8
4. If estimates came from a subgroup analysis, were the groups prespecified at the beginning of the study?	1	1
5. Was uncertainty handled by (1) statistical analysis to address random events, (2) sensitivity analysis to cover a range of assumptions?	9	5
6. Was incremental analysis performed between alternatives for resources and costs?	6	3
7. Was the methodology for data abstraction (including the value of health states and other benefits) stated?	5	5
8. Did the analytic horizon allow time for all relevant and important outcomes? Were benefits and costs that went beyond 1 year discounted (3% to 5%) and justification given for the discount rate?	7	7
9. Was the measurement of costs appropriate and the methodology for the estimation of quantities and unit costs clearly described?	8	8
10. Were the primary outcome measure(s) for the economic evaluation clearly stated and did they include the major short-term, long-term and negative outcomes included?	6	5
11. Were the health outcomes measures/scales valid and reliable? If previously tested valid and reliable measures were not available, was justification given for the measures/scales used?	7	7
12. Were the economic model (including structure), study methods and analysis, and the components of the numerator and denominator displayed in a clear, transparent manner?	8	6
13. Were the choice of economic model, main assumptions, and limitations of the study stated and justified?	7	5
14. Did the author(s) explicitly discuss direction and magnitude of potential biases?	6	6
15. Were the conclusions/recommendations of the study justified and based on the study results?	8	8
16. Was there a statement disclosing the source of funding for the study?	3	3
TOTAL POINTS	**100**	**86**

QHES Instrument[9] **Study** Garrison

Questions	Possible Points	Points Awarded
1. Was the study objective presented in a clear, specific, and measurable manner?	7	5
2. Were the perspective of the analysis (societal, third-party payer, etc.) and reasons for its selection stated?	4	2
3. Were variable estimates used in the analysis from the best available source (ie, randomized controlled trial - best, expert opinion - worst)?	8	6
4. If estimates came from a subgroup analysis, were the groups prespecified at the beginning of the study?	1	1
5. Was uncertainty handled by (1) statistical analysis to address random events, (2) sensitivity analysis to cover a range of assumptions?	9	6
6. Was incremental analysis performed between alternatives for resources and costs?	6	6
7. Was the methodology for data abstraction (including the value of health states and other benefits) stated?	5	3
8. Did the analytic horizon allow time for all relevant and important outcomes? Were benefits and costs that went beyond 1 year discounted (3% to 5%) and justification given for the discount rate?	7	5
9. Was the measurement of costs appropriate and the methodology for the estimation of quantities and unit costs clearly described?	8	8
10. Were the primary outcome measure(s) for the economic evaluation clearly stated and did they include the major short-term, long-term and negative outcomes included?	6	5
11. Were the health outcomes measures/scales valid and reliable? If previously tested valid and reliable measures were not available, was justification given for the measures/scales used?	7	5
12. Were the economic model (including structure), study methods and analysis, and the components of the numerator and denominator displayed in a clear, transparent manner?	8	3
13. Were the choice of economic model, main assumptions, and limitations of the study stated and justified?	7	5
14. Did the author(s) explicitly discuss direction and magnitude of potential biases?	6	4
15. Were the conclusions/recommendations of the study justified and based on the study results?	8	8
16. Was there a statement disclosing the source of funding for the study?	3	0
TOTAL POINTS	**100**	**72**

QHES Instrument[9] **Study** Karppinen et al (2001)[11, 12]

Questions	Possible Points	Points Awarded
1. Was the study objective presented in a clear, specific, and measurable manner?	7	0
2. Were the perspective of the analysis (societal, third-party payer, etc.) and reasons for its selection stated?	4	0
3. Were variable estimates used in the analysis from the best available source (ie, randomized controlled trial - best, expert opinion - worst)?	8	8
4. If estimates came from a subgroup analysis, were the groups prespecified at the beginning of the study?	1	0
5. Was uncertainty handled by (1) statistical analysis to address random events, (2) sensitivity analysis to cover a range of assumptions?	9	0
6. Was incremental analysis performed between alternatives for resources and costs?	6	6
7. Was the methodology for data abstraction (including the value of health states and other benefits) stated?	5	5
8. Did the analytic horizon allow time for all relevant and important outcomes? Were benefits and costs that went beyond 1 year discounted (3% to 5%) and justification given for the discount rate?	7	0
9. Was the measurement of costs appropriate and the methodology for the estimation of quantities and unit costs clearly described?	8	0
10. Were the primary outcome measure(s) for the economic evaluation clearly stated and did they include the major short-term, long-term and negative outcomes included?	6	6
11. Were the health outcomes measures/scales valid and reliable? If previously tested valid and reliable measures were not available, was justification given for the measures/scales used?	7	7
12. Were the economic model (including structure), study methods and analysis, and the components of the numerator and denominator displayed in a clear, transparent manner?	8	8
13. Were the choice of economic model, main assumptions, and limitations of the study stated and justified?	7	0
14. Did the author(s) explicitly discuss direction and magnitude of potential biases?	6	6
15. Were the conclusions/recommendations of the study justified and based on the study results?	8	0
16. Was there a statement disclosing the source of funding for the study?	3	3
TOTAL POINTS	**100**	**49**

Appendix E. LEVEL OF EVIDENCE FOR COMPARATIVE STUDIES

Methodological quality of therapeutic studies evaluating the efficacy or effectiveness of rhBMP-2 on-label use in the lumbar spine

Methodological principle	Boden 2000	Burkus 2002	Burkus 2003
Study Design			
Randomized controlled trial	√	√	
→ Random sequence generation*	-	-	
→ Allocation concealment*	-	-	
→ Intention to treat*	+/-	+/-	
Cohort study			√ (integrated analysis)
Case-control study			
Case series			
Other Methods Implementation			
Independent or blind assessment	+	+	+
Co-interventions applied equally	+	+	+
Complete follow-up of ≥ 80%	+	+	+
Adequate sample size	-	+/-	+
Controlling for possible confounding†	-	+	+
Evidence class	**IIb**	**IIb**	**II**

* Applies to randomized controlled trials only.
† Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.

Methodological quality of therapeutic studies evaluating the efficacy or effectiveness of rhBMP-2 off-label use in the lumbar spine

Methodological principle	Boden 2002	Burkus 2005/ 2006	Dawson 2009	Dimar 2009
Study Design				
Randomized controlled trial	√	√	√	√
→ Random sequence generation*	-	+	+	+
→ Allocation concealment*	-	+	-	-
→ Intention to treat*	+/-	+/-	+/-	-
Cohort study				
Case-control study				
Case series				
Other Methods Implementation				
Independent or blind assessment	+	+	+	+
Co-interventions applied equally	+	+	+	+
Complete follow-up of ≥ 80%	+	+	+	+
Adequate sample size	+/-	+	+	+
Controlling for possible confounding†	-	-	-	+/-
Evidence class	**IIb**	**IIa**	**IIb**	**IIb**

Methodological principle	Glassman 2008	Haid 2004	Glassman 2007	Mummaneni 2004
Study Design				
Randomized controlled trial	√	√		
→ Random sequence generation*	-	-		
→ Allocation concealment*	-	-		
→ Intention to treat*	+/-	+/-		
Cohort study			√	√
Case-control study				
Case series				
Other Methods Implementation				
Independent or blind assessment	+	+	+	+/-
Co-interventions applied equally	-	+	-	+
Complete follow-up of ≥ 80%	+	+	-	+
Adequate sample size	+	+	-	-
Controlling for possible confounding†	-	+	+/-	-
Evidence class	**IIb**	**IIb**	**III**	**III**

Continued on next page…

Methodological principle	Pradhan 2006	Singh 2006	Slosar 2007	Glassman Dimar 2007‡
Study Design				
Randomized controlled trial				
→ Random sequence generation*				
→ Allocation concealment*				
→ Intention to treat*				
Cohort study	√		√	√
Case-control study		√		
Case series				
Other Methods Implementation				
Independent or blind assessment	+	+	+	+
Co-interventions applied equally	+/-	+/-	+	+
Complete follow-up of ≥ 80%	+	+	+	-
Adequate sample size	-	-	+	+
Controlling for possible confounding†	-	-	+	+
Evidence class	**III**	**III**	**III**	**III**

Methodological principle	Carragee 2011	Crawford 2010	Howard 2011	Joesph 2007
Study Design				
Randomized controlled trial				
→ Random sequence generation*				
→ Allocation concealment*				
→ Intention to treat*				
Cohort study	√	√	√ (cross-sectional)	√
Case-control study				
Case series				
Other Methods Implementation				
Independent or blind assessment	+	+	+	+
Co-interventions applied equally	+	-	-	-
Complete follow-up of ≥ 80%	+	+	-	+
Adequate sample size	+	-	-	+
Controlling for possible confounding†	+	-	-	-
Evidence class	**II**	**III**	**III**	**III**

Continued on next page…

Methodological principle	Latzman 2010	Lee 2010	Rihn 2009	Taghavi 2010
Study Design				
Randomized controlled trial				
→ Random sequence generation*				
→ Allocation concealment*				
→ Intention to treat*				
Cohort study	√	√	√	√
Case-control study				
Case series				
Other Methods Implementation				
Independent or blind assessment	-	+	-	+
Co-interventions applied equally	-	-	-	+
Complete follow-up of ≥ 80%	-	-	+	-
Adequate sample size	+	+	-	+
Controlling for possible confounding†	+	-	-	+
Evidence class	**III**	**III**	**III**	**III**

Methodological principle	Vaidya, Weir (2007)	Burkus 2011
Study Design		
Randomized controlled trial		
→ Random sequence generation*		
→ Allocation concealment*		
→ Intention to treat*		
Cohort study	√	√ (integrated analysis)
Case-control study		
Case series		
Other Methods Implementation		
Independent or blind assessment	+	-
Co-interventions applied equally	+	-
Complete follow-up of ≥ 80%	+	-
Adequate sample size	+	+
Controlling for possible confounding†	-	-
Evidence class	**III**	**III**

* Applies to randomized controlled trials only.
† Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.
‡ Differential efficacy subset analysis of Dimar (2009) RCT.

Methodological quality of therapeutic studies evaluating the efficacy or effectiveness of rhBMP-7 off-label use in the lumbar spine

Methodological principle	Johnsson 2002	Kanayama 2006	Vaccaro, Lawrence 2008/ Hwang 2010
Study Design			
Randomized controlled trial	√	√	√
→ Random sequence generation*	+	-	+
→ Allocation concealment*	+/-	-	-
→ Intention to treat*	+/-	+/-	+/-
Cohort study			
Case-control study			
Case series			
Other Methods Implementation			
Independent or blind assessment	+	+/-	+
Co-interventions applied equally	+	+	+
Complete follow-up of ≥ 80%	+	+	-
Adequate sample size	-	-	+/-
Controlling for possible confounding†	-	-	+/-
Evidence class	**IIb**	**IIb**	**IIb**

Methodological principle	Vaccaro 2004/2005/ 2008	Delawi 2010
Study Design		
Randomized controlled trial	√	√
→ Random sequence generation*	+	+
→ Allocation concealment*	-	+/-
→ Intention to treat*	+/-	-
Cohort study		
Case-control study		
Case series		
Other Methods Implementation		
Independent or blind assessment	+	+
Co-interventions applied equally	+	+
Complete follow-up of ≥ 80%	-	+
Adequate sample size	-	-
Controlling for possible confounding†	+	+/-
Evidence class	**IIb**	**IIb**

* Applies to randomized controlled trials only.
† Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.

Methodological quality of therapeutic studies evaluating the efficacy or effectiveness of any BMP use in the lumbar, cervical, and/or thoracic spine

Methodological principle	Cahill 2009	Cahill 2011	Deyo 2011
Study Design			
Randomized controlled trial			
→ Random sequence generation*			
→ Allocation concealment*			
→ Intention to treat*			
Cohort study	√ (database)		√ (database)
Case-control study		√ (database)	
Case series			
Other Methods Implementation			
Independent or blind assessment	-	+	+
Co-interventions applied equally	+	+	+
Complete follow-up of ≥ 80%	-	-	-
Adequate sample size	+	+	+
Controlling for possible confounding†	+	+	+
Evidence class	**III**	**III**	**III**

Methodological principle	Mines 2011	Williams 2011
Study Design		
Randomized controlled trial		
→ Random sequence generation*		
→ Allocation concealment*		
→ Intention to treat*		
Cohort study	√ (database)	
Case-control study		
Case series		
Other Methods Implementation		
Independent or blind assessment	+	-
Co-interventions applied equally	-	-
Complete follow-up of ≥ 80%	+	-
Adequate sample size	+	+
Controlling for possible confounding†	+	-
Evidence class	**III**	**III**

* Applies to randomized controlled trials only.
† Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.

Methodological quality of therapeutic studies evaluating the efficacy or effectiveness of rhBMP-2 off-label use in the cervical spine

Methodological principle	Baskin 2003	Buttermann 2008	Crawford 2009	Smucker 2006
Study Design				
Randomized controlled trial	√			
→ Random sequence generation*	-			
→ Allocation concealment*	-			
→ Intention to treat*	+/-			
Cohort study		√	√	
Case-control study				√
Case series				
Other Methods Implementation				
Independent or blind assessment	+	+	-	-
Co-interventions applied equally	+	-	+/-	-
Complete follow-up of ≥ 80%	+	+	+	+
Adequate sample size	+	+	-	+
Controlling for possible confounding†	-	-	+	+
Evidence class	**IIb**	**III**	**III**	**III**

Methodological principle	Vaidya, Carp 2007	Vaidya, Weir 2007	Xu 2011	Yaremchuk 2010
Study Design				
Randomized controlled trial				
→ Random sequence generation*				
→ Allocation concealment*				
→ Intention to treat*				
Cohort study	√	√	√	√
Case-control study				
Case series				
Other Methods Implementation				
Independent or blind assessment	+	+	-	-
Co-interventions applied equally	+	+	-	-
Complete follow-up of ≥ 80%	+	+	+	-
Adequate sample size	+	+	+	+
Controlling for possible confounding†	-	-	+	-
Evidence class	**III**	**III**	**III**	**III**

* Applies to randomized controlled trials only.
† Authors must provide a description of robust baseline characteristics, and control for those that are unequally distributed between treatment groups.

Notes on LoE grades: comparative studies
Partial credit given for:

- Intent to treat: no explicit statement OR patients classified as failures were excluded after reoperation. (If the patients classified as failures were excluded entirely, no credit given)
- Adequate sample size: Large study > 200 pts but not statistically meaningful differences b/w groups

BMP2 on-label (lumbar)

1. Boden 2000[13] (AHRQ ref 71)
    a. Study design:
        i. Random sequence generation: "marginal balancing method"- no credit; not adequately described (AHRQ agrees). (authors reference a book)
        ii. Allocation concealment: no credit; no information
        iii. Intention to treat: partial credit; no explicit statement but data appear to have been handled this way
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported.
    c. Co-interventions applied equally: credit given; no obvious discrepancies
    d. Complete f/u of ≥ 80%: yes (100%)
    e. Adequate sample size: no credit; only 14 pts enrolled (3 in control group)
    f. Controlling for possible confounding: no credit given; there were differences in pt weight between groups and no multivariate analysis was done.

2. Burkus 2002[14] (AHRQ ref 72)
    a. Study design:
        i. Random sequence generation: no credit; no information
        ii. Allocation concealment: no credit; no information
        iii. Intention to treat: partial credit; no explicit statement, and it appears that data from patients classified as failures (ie., had to undergo device removals, revisions, or supplemental fixations) were not reported after they had failed the treatment.
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported.
    c. Co-interventions applied equally: credit given; no obvious discrepancies
    d. Complete f/u of ≥ 80%: yes, 83% (232/279) at 24 mos.
    e. Adequate sample size: partial credit given, this study had a large number of patients enrolled (N = 279), however, there were no meaningful differences in outcomes between groups.
    f. Controlling for possible confounding: credit given; adequate table 1 & similar baseline scores b/w groups.

3. Burkus 2003[15] (AHRQ ref 182)
    a. Study design: retrospective integrated analysis of comparative data (cohort study)
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) (not stated explicitly but "all of the 679 pts were included in… studies using the same outcome measurement tools and methodology of analysis", thus can use Burkus 2002 as an example of how radiographs were assessed) or patient-reported.
    c. Co-interventions applied equally: credit given; no obvious discrepancies
    d. Complete f/u of ≥ 80%: yes (85%)
    e. Adequate sample size: yes, there were statistically meaningful differences in outcomes (ex: 24 mos radiographic success rates), also N = 679
    f. Controlling for possible confounding: credit given, "among 20 summarized variables, seven were found to be significantly different b/w the combined INFUSE gp and the combined autograft group;" these prognostic factors were then controlled for using analysis of covariance.

BMP7 on-label (lumbar)

No comparative studies

BMP2 off-label (lumbar)

1. Boden 2002[16] (AHRQ ref 84)
    a. Study design: RCT
        i. Random sequence generation: no credit; no information
        ii. Allocation concealment: no credit; no information
        iii. Intention to treat: partial credit; no explicit statement but data appear to have been handled this way (no mention of failures)
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported.
    c. Co-interventions applied equally: credit given; no obvious discrepancies
    d. Complete f/u of ≥ 80%: yes (93%)
    e. Adequate sample size: partial credit given; while there were statistically meaningful differences b/w groups in the fusion rates, there were only 5 pts in the control group and 9 pts in the BMP-2 only group.
    f. Controlling for possible confounding: no credit; 40% of patients in the autograft group had diabetes compared with 0% in either treatment group (P = .036); differences not controlled for.

2. Burkus 2005[17] (AHRQ ref 85) (includes Burkus 2006[18] safety data only from same study)
    a. Study design: RCT

i. Random sequence generation: credit given; statistical program (SAS) used to produce sequentially numbered envelopes specific to each enrollment site.
ii. Allocation concealment: credit given; surgeons blinded to randomization schedule; allocation in sequentially numbered envelopes.
iii. Intention to treat: partial credit; no explicit statement, and data from patients classified as failures (ie., had to undergo device removals, revisions, or supplemental fixations) were not reported after they had failed the treatment.

b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported.
c. Co-interventions applied equally: credit given; no obvious discrepancies
d. Complete f/u of ≥ 80%: yes (96%)
e. Adequate sample size: credit given; there were statistically meaningful differences in several outcome measures
f. Controlling for possible confounding: no credit; there were statistically meaningful differences in preop back pain scores b/w groups (*P* = .039); this difference was not accounted for.

3. Dawson 2009[19] (AHRQ ref 73)
   a. Study design: RCT
      i. Random sequence generation: credit given; randomization stratified by site with a fixed block size of four.
      ii. Allocation concealment: no credit; no information given. After consent and randomization, two patients in each group elected not to participate in the study (whether pts were aware of their tx allocation was NR).
      iii. Intention to treat: partial credit given; a modified intent-to-treat principle was used in which patients who had failed had last available data carried forward.
   b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported.
   c. Co-interventions applied equally: credit given; no obvious discrepancies
   d. Complete f/u of ≥ 80%: yes (87% (40/46)) (even if accounting for the 4 patients randomized who then dropped out, complete f/u would be 80% (40/50)).
   e. Adequate sample size: credit given; there was a statistically meaningful difference in fusion rates at 6 mos between the BMP and control groups (91% vs 58%) (*P* = .032). Even at 24 mos, while the difference wasn't statistically meaningful, there was quite a difference in fusion rates b/w groups (95% vs 67%).
   f. Controlling for possible confounding: no credit; baseline scores for ODI, pain, etc. were not reported. Even though the authors reported mean

change score, we don't know if the groups were comparable in baseline scores (which could potentially affect the outcomes). While the authors used regression analysis to control for differences in demographics (Workers' Comp, litigation, previous spinal surgery), this does not appear to have been done to control for potential differences in preop scores.

4. Dimar 2009[20] (AHRQ ref 86)
    a. Study design: RCT
        i. Random sequence generation: credit, "randomization was centrally generated on a 1:1 basis, stratified by site with use of a fixed block size of 4 and sealed envelopes with sequential numbers."
        ii. Allocation concealment: partial credit; randomization done off-site but there was no mention of opaque envelopes.
        iii. Intention to treat: no credit: use an as-treated analysis: "A small number of patients required an additional surgical procedure; their outcomes were recorded as a treatment failure. For other outcome variables, the last observations made before the additional surgical procedures or interventions were carrier forward with use of the last observation carrier forward technique for all future evaluation periods…The protocol predefined the as-treated analysis as the primary analysis for the study, on the basis of the statistical consideration that intent-to-treat analysis may not be conservative for assessing a noninferiorty hypothesis."… "There were two crossovers in the study. They were analyzed on the basis of the treatment received (the so-called as-treated analysis).
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported.
    c. Co-interventions applied equally: credit given; no obvious discrepancies
    d. Complete f/u of $\geq$ 80%: yes (89%)
    e. Adequate sample size: yes, differences in some outcomes (operative time, blood loss, adverse events), and n > 100 for both groups.
    f. Controlling for possible confounding: partial credit, there was only one statistically meaningful baseline difference (spinal litigation) in the 15+ baseline characteristics reported or baseline data, but it wasn't controlled for or discussed.

5. Glassman 2008[21] (AHRQ ref 87)
    a. Study design: RCT
        i. Random sequence generation: no credit; no information provided on method of randomization
        ii. Allocation concealment: no credit; no information provided on method of randomization
        iii. Intention to treat: partial credit given; ITT explicitly stated ("One pt in ICBG group ended up receiving BMP but was analyzed as part of the ICBG group in an intent to treat analysis.")

BUT patients who failed treatment and required revision procedure had last observation carried forward.

b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported.
c. Co-interventions applied equally: no credit given; while the mean # of levels fused was similar b/w groups, the authors did not report the # of pts in each group that underwent 1-, 2-, 3-, etc.- level fusion
d. Complete f/u of ≥ 80%: yes (94%)
e. Adequate sample size: yes, differences in some outcomes (frequency of leg pain, operative time, # periop complications)
f. Controlling for possible confounding: no credit given; there were no differences in baseline characteristics, but there was a difference in preoperative leg pain scores between groups that was not controlled for.

6. Haid 2004[22] (AHRQ ref 88)
    a. Study design: RCT
        i. Random sequence generation: no credit; no information provided on method of randomization
        ii. Allocation concealment: no credit; no information provided on method of randomization
        iii. Intention to treat: partial credit; no explicit statement but data appear to have been handled this way (no mention of failures being withdrawn from analysis)
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported.
    c. Co-interventions applied equally: credit given; no obvious discrepancies
    d. Complete f/u of ≥ 80%: yes (94%)
    e. Adequate sample size: yes, statistically meaningful difference in back pain b/w groups at 24 mos.
    f. Controlling for possible confounding: credit given, similar baseline characteristics/preop scores.

7. Glassman 2007[23] (AHRQ ref 99)
    a. Study design: retrospective cohort with historical control
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; radiographic outcomes (which were the only outcomes) were evaluated by independent surgeons (for both groups).
    c. Co-interventions applied equally: no credit given; BMP group received supplemental bone graft extenders or fillers; insufficient detail for (historical) control group or whether these patients also received bone graft extenders or fillers; no info on # of patients in the control group who received 1-, 2- or 3-level fusion.

d. Complete f/u of ≥ 80%: no credit; f/u NR. Retrospective cohort; only patients with 2-yr f/u were included, which eliminates patients who may have shorter f/u etc.
e. Adequate sample size: no credit; no comparative data presented to give idea of whether there were statistically meaningful differences b/w groups.
f. Controlling for possible confounding: partial credit given; the authors noted "there were no statistically significant demographic differences between the [groups]", however, preoperative diagnosis was only reported for the BMP group.

8. Mummaneni 2004[24] (AHRQ ref 100)
    a. Study design: retrospective cohort study
    b. Independent or blind assessment: partial credit; no mention that radiographs were evaluated in a blinded or independent manner; but VAS and Prolo scales are patient-reported
    c. Co-interventions applied equally: credit given; no obvious discrepancies (all but 2 pts underwent single-level fusion)
    d. Complete f/u of ≥ 80%: yes (91%)
    e. Adequate sample size: no; all results similar b/w groups; N < 50
    f. Controlling for possible confounding: no credit; there was a statistically meaningful difference in the % of pts over the age of 65 b/w groups (24% vs 0%, $P < .01$) that was not controlled for.

9. Pradhan 2006[25] (AHRQ ref 101)
    a. Study design: prospective cohort with historical control
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; radiographic outcomes (which were the only outcomes) were evaluated by an independent/blinded surgeon.
    c. Co-interventions applied equally: no credit given; no obvious discrepancies except that the control group was a historical control, and no dates were provided for surgery.
    d. Complete f/u of ≥ 80%: yes (100%)
    e. Adequate sample size: no credit; BMP group had only 9 pts
    f. Controlling for possible confounding: no; differences in gender b/w groups (33% vs 23% males), % of pts who smoke was NR; also differences in length f/u (36 vs 26 mos) that could affect results since results are presented for final f/u.

10. Singh 2006[26] (AHRQ ref 102)
    a. Study design: prospective case control study
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; radiographic outcomes (which were the only outcomes) were evaluated by an independent/blinded surgeon.
    c. Co-interventions applied equally: no credit given; while the mean # of levels fused was similar b/w groups (1.79 vs 2.0), the authors did not report the # of pts in each group that underwent 1-, 2-, 3-, etc.- level fusion

d. Complete f/u of ≥ 80%: yes (96%)
e. Adequate sample size: no credit; control group had 11 patients (vs 39 in the BMP group)
f. Controlling for possible confounding: no credit; ICBG group was a mean of 11 years younger (54 years vs 65 years), which was not controlled for.

11. Slosar 2007[27] (AHRQ ref 103)
    a. Study design: prospective cohort
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported
    c. Co-interventions applied equally: credit given; no obvious discrepancies (the numbers of patients who underwent 1-, 2-, and 3-level fusion were similar between groups)
    d. Complete f/u of ≥ 80%: yes (96%)
    e. Adequate sample size: yes; statistically meaningful difference in fusion rates b/w groups
    f. Controlling for possible confounding: yes; table 1 demonstrates similar baseline characteristics b/w groups.

12. Glassman Dimar 2007[28]
    a. Study design: retrospective analysis of subset of patient in Dimar 2009 RCT
    b. Independent or blind assessment: Credit given; assessment of all radiographic parameters done by independent radiologists who were blinded to treatment group; all other outcomes were patient-reported.
    c. Co-interventions applied equally: credit given; no obvious discrepancies between groups.
    d. Complete f/u of ≥ 80%: no credit, % f/u not reported: no explicit stmt about LTF, but successful fusion rates given seem to suggest all subjects available at 2 years. However, study could have selected only patients with f/u at 2 years for this analysis, can't tell from article.
    e. Adequate sample size: credit given; there were stat sig differences in fusion rate between smokers versus non-smokers at 24 mos. (85.7% vs 97.2%) ($P = .016$).
    f. Controlling for possible confounding: credit given; "there were no statistically significant differences in demographic parameters between four smoking/graft montage subgroups".

13. Carragee 2011[29]
    a. Study design: retrospective cohort study
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; "postoperative outcomes were recorded by independent research assistants in a deidentified database," all primary outcomes were patient-reported (RE only)

c. Co-interventions applied equally: credit given; no obvious discrepancies
d. Complete f/u of ≥ 80%: yes, 100%
e. Adequate sample size: yes, statistically more RE events in BMP group vs. control group
f. Controlling for possible confounding: credit given, similar baseline data b/w groups.

14. Crawford 2010[30]
    a. Study design: retrospective cohort study
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported
    c. Co-interventions applied equally: no credit given; control group operated before BMP available (1998-2002) while BMP group underwent surgery b/w 2002-2006; differences in # of anterior levels fused and approach used.
    d. Complete f/u of ≥ 80%: yes (94%)
    e. Adequate sample size: no; no outcomes of clinical significance had statistically meaningful differences b/w groups.
    f. Controlling for possible confounding: no credit; statistically meaningful differences in age, baseline mental health SRS scores, and length f/u b/w groups that were not controlled for.

15. Howard 2011[31]
    a. Study design: cross-sectional study (treated as cohort study for LoE grading)
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; the primary outcome of graft site pain was reported by patient upon exam by an independent and blinded investigator (no scar over graft site in these patients as the graft was harvested through the midline lumbar incision)
    c. Co-interventions applied equally: no credit given, very little info was given on interventions, including how fusion was done, when surgery performed (were the BMP and control pts seen around the same time?), mean # of levels fused per group (pts had 1- or 2-level fusion).
    d. Complete f/u of ≥ 80%: no, f/u NR
    e. Adequate sample size: no, no differences in outcomes measured b/w tx groups
    f. Controlling for possible confounding: no credit given, baseline data NR separately for each tx group.

16. Joseph 2007[32]
    a. Study design: prospective cohort study
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; the primary outcome (radiographic: heterotopic bone

formation, fusion) was independently reviewed (study described as an "independent CT analysis").
c. Co-interventions applied equally: no credit; overall number of pts with 1- and 2-level fusions given, but not given per treatment group.
d. Complete f/u of ≥ 80%: yes (33/34)
e. Adequate sample size: credit given, there was a statistically meaningful difference in fusion rates at 6 mos.
f. Controlling for possible confounding: no credit given, baseline data NR separately for each tx group.

17. Latzman 2010[33]
a. Study design: retrospective cohort study
b. Independent or blind assessment: no credit given; retrospective study and no mention that lab tests done to determine creatine and BUN levels (primary outcomes) were done in an independent or blind fashion.
c. Co-interventions applied equally: no credit given, very little info was given on interventions, including how fusion was done, when surgery performed (were the BMP and control pts seen around the same time?), and in addition, the control group had significantly longer mean f/u than the BMP group (4.49 ± 2.0 vs. 1.48 ± 0.85, $P < .001$). In addition, there were statistically more patients in the BMP group that received an interbody cage (70% vs 30%, $P = .001$).
d. Complete f/u of ≥ 80%: no, f/u NR
e. Adequate sample size: credit given, meaningful difference in the percentage of patients b/w groups who had transient renal failure.
f. Controlling for possible confounding: credit given; similar baseline demographics and preoperative BUN, creatine levels between groups.

18. Lee 2010[34]
a. Study design: retrospective cohort study
b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported
c. Co-interventions applied equally: no credit given; there were differences in the percentage of patients between groups who were undergoing primary vs revision surgery (BMP 65+ year vs. BMP <65 years vs. IBCG) (35% vs 50% vs 20%) and the patients were undergoing single or multilevel fusion (50% vs. 25% vs. 68% of patients underwent multilevel fusions)
d. Complete f/u of ≥ 80%: no credit; f/u NR
e. Adequate sample size: credit given, statistically meaningful differences b/w groups at 2 yrs in pain
f. Controlling for possible confounding: credit; differences in baseline demographics (sex, presence of comorbidities, presence of osteoporosis, smoking status, primary vs revision surgery), BUT these were controlled for by multivariable analyses.

19. Rihn 2009[35]
    a. Study design: retrospective cohort study
    b. Independent or blind assessment: no credit given; blinding of surgeons/pts not possible; no info on blinding of assessors (for complications), the majority of which aren't patient reported. Because this is retrospective blinding is unlikely.
    c. Co-interventions applied equally: no credit given; patients in the autograft group had significantly longer mean length follow-up compared with the patients in the BMP group (35.8 vs. 24.4 mos., respectively; $P < .001$).
    d. Complete f/u of ≥ 80%: yes (91%)
    e. Adequate sample size: no credit given; there no meaningful differences b/w groups
    f. Controlling for possible confounding: no credit given; demographics NR separately for each treatment group.

20. Taghavi 2010[36]
    a. Study design: retrospective cohort study
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported
    c. Co-interventions applied equally: credit given; no obvious discrepancies (while there was no difference in number of levels per patient, there were differences in the percentages of patients undergoing single vs multilevel fusion b/w groups (rhBMP2 vs BMAA vs autograft) (54% vs 39% vs 50% undergoing single-level fusion, HOWEVER these were controlled for by stratified analyses)
    d. Complete f/u of ≥ 80%: no credit; f/u NR (only pts with minimum 2 year f/u were included)
    e. Adequate sample size: credit given; there were differences in time to solid fusion b/w groups
    f. Controlling for possible confounding: credit given; similar baseline demographics and VAS pain scores.

21. Vaidya, Weir 2007[37]
    a. Study design: prospective cohort study
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; radiographs assessed by two independent observers, no other outcomes
    c. Co-interventions applied equally: no credit given; while the mean # of levels fused was similar b/w groups, the authors did not report the # of pts in each group that underwent 1-, 2-, 3-, etc.- level fusion.
    d. Complete f/u of ≥ 80%: yes (100%)
    e. Adequate sample size: credit given; there was a meaningful differences b/w groups in the mean subsidence between groups for TLIF pts

f. Controlling for possible confounding: no credit given; very limited demographics reported, demographics NR separately for lumbar vs cervical pt, no info on comorbidities,.

22. Burkus 2011[38]
    a. Study design: integrated analysis of 3 previous studies (treat as cohort study)
    b. Independent or blind assessment: No credit given; blinding of surgeons/pts not possible; no info given on blinding of investigators performing the antibody tests.
    c. Co-interventions applied equally: no credit given; very little info given, including dose of rhBMP2 (and one of the studies has been published in abstract form only, so details unavailable for those patients).
    d. Complete f/u of $\geq$ 80%: credit; f/u NR (only pts with minimum 2 year f/u were included)
    e. Adequate sample size: credit given; (N = 1493)
    f. Controlling for possible confounding: no credit given; very little info given (and one of the studies has been published in abstract form only, so details unavailable for those patients).

BMP7 off-label (lumbar)

1. Vaccaro pilot study 2004[39]/2005[40]/2008[41] (AHRQ refs 184, 185, 95)
    a. Study design: RCT
        i. Random sequence generation: credit; "the randomization allocation was performed in SAS using the PLAN procedure."
        ii. Allocation concealment: no credit; no mention of concealment; "A designated representative from the study sponsor informed the site as to which treatment group the subject was to be enrolled in before the time of his/her spinal fusion." No further details provided.
        iii. Intention to treat: partial credit; no explicit statement; last-observation carried forward data provided separately.
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported
    c. Co-interventions applied equally: credit given; no obvious discrepancies
    d. Complete f/u of ≥ 80%: no, radiographic results (58%), clinical results (72%)
    e. Adequate sample size: no, control group n = 12; low % f/u combined with small patient enrollment makes it difficult to determine whether any differences were meaningful.
    f. Controlling for possible confounding: partial credit, similar demographics and baseline scores EXCEPT presence of straight leg tension sign causing leg pain at baseline (OP1 vs autograft) (29% vs 0%) that was not controlled for.

2. Johnsson 2002[42] (AHRQ ref 92)
    a. Study design: RCT
        i. Random sequence generation: credit; randomization performed in blocks of six patients
        ii. Allocation concealment: partial credit given; patient and surgeon blinded until procedure began, but no information was provided as to how concealment was ensured.
        iii. Intention to treat: partial credit; no explicit statement
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported
    c. Co-interventions applied equally: credit given; no obvious discrepancies
    d. Complete f/u of ≥ 80%: yes (100%)
    e. Adequate sample size: no, only 10 pts per tx group
    f. Controlling for possible confounding: no credit; poorly described demographics; OP-1 group had 30% males while control group had 50% males which was not controlled for

3. Kanayama 2006[43] (AHRQ ref 93)
    a. Study design: RCT

i. Random sequence generation: no credit; details NR
ii. Allocation concealment: no credit; details NR
iii. Intention to treat: partial credit; no explicit statement

b. Independent or blind assessment: Partial credit given; blinding of surgeons/pts not possible; no reporting that the radiologist were blinded or independent but the other primary outcome was patient-reported (ODI)
c. Co-interventions applied equally: credit given; no obvious discrepancies
d. Complete f/u of ≥ 80%: yes (95%)
e. Adequate sample size: no, 10 pts per tx group
f. Controlling for possible confounding: no, pts in BMP7 group were older than those in the control group (70 vs. 59 years, $P < .05$), which was not controlled for.

4. Vaccaro, Lawrence 2008[44] (AHRQ ref 94); Hwang 2010[45] (only additional relevant info was deaths)
    a. Study design: RCT
        i. Random sequence generation: credit; "randomization was performed after enrollment but before surgery using a computerized algorithm (SAS)"
        ii. Allocation concealment: no credit given; "patients and physicians became aware of the treatment assignment at the time of the randomization and before surgery."
        iii. Intention to treat: partial credit given; a modified intent-to-treat principle was used in which patients who had failed (or died) were excluded from further analysis.
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported
    c. Co-interventions applied equally: credit given; no obvious discrepancies
    d. Complete f/u of ≥ 80%: no; (60%): 335 enrolled and randomized, 295 treated (40 patients either withdrew or were excluded based on the inclusion/exclusion criteria); at 36 + mos, 202 pts were evaluated (202/335)
    e. Adequate sample size: partial credit given; while there were no statistically meaningful differences between groups, the study was large (N = 335)
    f. Controlling for possible confounding: yes, similar demographics and baseline scores

5. Delawi 2010[46]
    a. Study design: RCT
        i. Random sequence generation: credit given; computer-generated randomization code produced according to the "random permuted block" by an independent researcher using SYSTAT
        ii. Allocation concealment: partial credit: no mention of opaque envelopes; "surgeons were blinded to the treatment group as long

as possible. That means that the decompression and placement of the screws were performed before the envelope containing the randomization of the patient was opened and the surgeon received the result of the randomization."

iii. Intention to treat: no credit; one patient in the autograft group received local autograft only (no ICBG), and the patient was excluded from analysis.

b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported
c. Co-interventions applied equally: credit given; no obvious discrepancies
d. Complete f/u of $\geq$ 80%: yes (89%)
e. Adequate sample size: no credit; there were no statistically meaningful differences in the outcomes b/w the groups, sample size small ($n \leq 16$ per group)
f. Controlling for possible confounding: partial credit; there were no statistically meaningful differences b/w groups at baseline for demographics or ODI scores, BUT the distribution of which spinal level fused was quite different b/w groups, and this was not controlled for

BMP (any) off-label (lumbar) (Database studies)

1. Cahill 2009[47]
    a. Study design: Retrospective cohort study (database study)
    b. Independent or blind assessment: No credit given; blinding of surgeons/pts not possible; database study with records reviewed retrospectively. No info on blinding of assessors (for complications), the majority of which aren't patient reported ("any complication", dysphagia or hoarseness, wound complication, "other complications"). Because this is retrospective blinding is unlikely.
    c. Co-interventions applied equally: credit given; no obvious discrepancies.
    d. Complete f/u of ≥ 80%: No, % f/u NR
    e. Adequate sample size: credit given; there were meaningful differences in the percentage of patients between groups with complications (anterior cervical pts)
    f. Controlling for possible confounding: credit given; extensive demographic info given; no statistically meaningful differences between groups. Multivariate analysis also done to adjust for significant predictors as well.

2. Cahill 2011[48]
    a. Study design: Case-control (retrospective) (database study)
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; database study with records reviewed retrospectively. Reliable data sources used: outcomes reported not subject to opinion (length of stay, readmission, repeat fusion).
    c. Co-interventions applied equally: credit given; no obvious discrepancies.
    d. Complete f/u of ≥ 80%: No, % f/u NR (all patients included had follow-up of at least 12 months)
    e. Adequate sample size: credit given; there were meaningful differences in the percentage of patients between groups requiring repeat fusion procedures.
    f. Controlling for possible confounding: credit given; extensive demographic info given; no statistically meaningful differences between groups. Multivariate analysis also done to adjust for significant predictors as well.

3. Deyo 2011[49]
    a. Study design: Retrospective cohort study (database study)
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; database study with records reviewed retrospectively. Reliable data sources used: majority of outcomes reported not subject to opinion (readmission, death, nursing home discharge, reoperation).
    c. Co-interventions applied equally: credit given; no obvious discrepancies.
    d. Complete f/u of ≥ 80%: No, % f/u NR
    e. Adequate sample size: credit given; large study (BMP group: n = 1703, control group n = 15,119); there were meaningful differences in the

percentage of patients between groups who were discharged into a nursing home.

f. Controlling for possible confounding: credit given; extensive demographic info given; statistically meaningful differences in age, # levels fused, fusion type, and history of spinal surgery; but regression analysis done (the authors stated the differences between groups remained small and nonsignificant but data NR for the regression analysis).

4. Mines 2011[50]
    a. Study design: Retrospective cohort study (database study)
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; database study with records reviewed retrospectively. Reliable data sources used: outcomes reported not subject to opinion (pancreatic cancer, death).
    c. Co-interventions applied equally: no credit given; no detail of co-interventions described (number of levels, surgical approach, etc).
    d. Complete f/u of ≥ 80%: Yes ("nearly all study participants survived to the end of the f/u period (BMP, 96.9%, non-BMP, 94.9%).
    e. Adequate sample size: credit given; there were meaningful differences in the percentage of patients between groups who were discharged into a nursing home.
    f. Controlling for possible confounding: credit given; extensive demographic info given; statistically meaningful differences in age, gender, race, diabetes, prior cholecystectomy; but multivariate regression analysis done.

5. Williams 2011[51]
    a. Study design: retrospective cohort study (database study)
    b. Independent or blind assessment: no credit; no information on assessments.
    c. Co-interventions applied equally: no credit; no information on number of levels fused.
    d. Complete f/u of ≥ 80%: no credit; no explicit f/u, but mention of long-term f/u not in database, and no method to determine the completeness of data submission to the db.
    e. Adequate sample size: credit given; large sample size and statistically significant differences between BMP and non-BMP groups, including higher overall complication rate for BMP subgroup versus non-BMP subgroup for adult scoliosis pts (13.8% vs 9.3%, $P < .001$).
    f. Controlling for possible confounding: no credit; many stat sig diff btw BMP and non-BMP groups, including age, diagnosis (degenerative spinal disorder, spondylolisthesis), and revision procedures.

BMP2 off-label (cervical)

1. Baskin 2003[52] (AHRQ ref 89)
    a. Study design: RCT
        i. Random sequence generation: no credit; no info provided on randomization
        ii. Allocation concealment: no credit; "after randomization, neither the surgeon nor the patient was blinded to the treatment."
        iii. Intention to treat: partial credit; no explicit statement
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported
    c. Co-interventions applied equally: credit given; no obvious discrepancies
    d. Complete f/u of ≥ 80%: yes (88%)
    e. Adequate sample size: credit given; there were statistically meaningful differences in the change NDI scores b/w groups at 24 mos (52.7 vs. 36.9) ($P < .03$)
    f. Controlling for possible confounding: no credit; 28% (5/18) of the BMP pts used tobacco vs 47% (7/15) of the control patients (at baseline), which was not controlled for.

2. Buttermann 2008[53] (AHRQ ref 104)
    a. Study design: prospective cohort study
    b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were patient-reported (fusion NR).
    c. Co-interventions applied equally: no credit; there were differences between the BMP and ICBG groups in the percentages of patients who underwent 3-levels ACDF procedures (33% vs. 6%) as well as 1-level ACDF procedures (13% vs. 42%).
    d. Complete f/u of ≥ 80%: yes (100%)
    e. Adequate sample size: credit given; there were meaningful differences between groups in the percentage of patients who experienced postoperative dysphagia (50%, BMP vs. 14%, ICBG).
    f. Controlling for possible confounding: no credit; there were differences between the BMP and ICBG groups in the percentages of males (50% vs. 67%), smokers (37% vs. 53%). While the authors wrote "smoking status was unrelated to outcomes scores," no data were provided. No mention was made of whether there were differences in tx outcomes when stratified by patient sex.

3. Crawford 2009[54] (AHRQ ref 105)
    a. Study design: retrospective cohort study
    b. Independent or blind assessment: no credit given; blinding of surgeons/pts not possible; none of the outcomes were patient-reported (fusion NR).
    c. Co-interventions applied equally: no credit; while the mean # of levels fused was similar b/w groups, the authors did not report the # of pts in

each group that underwent 1-, 2-, 3-, etc.- level fusion; bone graft extenders used at surgeon's discretion but their use was NR.
   d. Complete f/u of ≥ 80%: yes (100%)
   e. Adequate sample size: no credit; no meaningful differences b/w groups in outcomes
   f. Controlling for possible confounding: credit given; baseline characteristics were similar b/w groups.

4. Smucker 2006[55] (AHRQ ref 106)
   a. Study design: retrospective cohort study
   b. Independent or blind assessment: no credit; only data reported were swelling complications and it wasn't clear whether these were assessed in a blind/independent manner (likely not as the study is a retrospective chart review).
   c. Co-interventions applied equally: no credit; the BMP had a higher average # of levels fused (2.2 vs 1.7, $P = .001$) and a higher % of pts having 3+ levels fused (44% vs. 27%, $P = .02$), less likely to have supplemental plate fixation (88% vs. 97%, $P = .02$), greater use of allograft (88% vs. 81%, $P < .001$).
   d. Complete f/u of ≥ 80%: yes, 100% of consecutive pts
   e. Adequate sample size: credit given; there were statistically meaningful differences b/w groups in cervical swelling complications.
   f. Controlling for possible confounding: credit given; similar baseline characteristics between groups.

5. Vaidya, Carp 2007[56] (AHRQ ref 107)
   a. Study design: retrospective cohort study
   b. Independent or blind assessment: Credit given; blinding of surgeons/pts not possible; all primary outcomes were blinded (radiologist) or patient-reported
   c. Co-interventions applied equally: credit given: no obvious discrepancies;; (there were no meaningful differences in the number of levels fused between the BMP and control groups: (1-level: 46% vs. 36%; 2-level: 41% vs. 42%; 3-level: 18% vs. 13%)
   d. Complete f/u of ≥ 80%: no (79% (46/58))
   e. Adequate sample size: credit given; there was a meaningful difference in the % of patients with postoperative dysphagia between groups ($P \leq .02$)
   f. Controlling for possible confounding: no credit, robust baseline characteristics not described (ie., only age, sex, and diagnosis were reported); there was also a potentially significant difference in the percentage of males b/w groups (32% vs. 45%) that was not controlled for.

6. Vaidya, Weir 2007[37] (see BMP-2 off-label lumbar section)

7. Xu 2011[57]
   a. Study design: retrospective cohort study

b. Independent or blind assessment: Credit not given; blinding of surgeons/pts not possible; no info on blinded/ independent analysis of radiographic outcomes
c. Co-interventions applied equally: no credit; while the mean # of levels fused was similar b/w groups, the authors did not report the # of pts in each group that underwent 1-, 2-, 3-, etc.- level fusion
d. Complete f/u of ≥ 80%: yes (83%)
e. Adequate sample size: credit given; there was a meaningful difference in the % of patients b/w groups with fusion, recurrent neck pain.
f. Controlling for possible confounding: credit given; statistically similar baseline characteristics between groups.

8. Yaremchuk 2010[58]
    a. Study design: retrospective (database) cohort study
    b. Independent or blind assessment: Credit not given; blinding of surgeons/pts not possible; no info on blinded/ independent analysis of radiographic outcomes
    c. Co-interventions applied equally: no credit; very little info reported, including the # of pts in each group that underwent 1-, 2-, 3-, etc.- level fusion
    d. Complete f/u of ≥ 80%: no (% f/u NR)
    e. Adequate sample size: credit given; there was a meaningful difference in the length of stay after surgery between groups, as well as the percentage of pts who had tracheotomies, unplanned intubations, readmission, dysphagia, dyspnea, and respiratory failure.
    f. Controlling for possible confounding: no credit given; no demographic info reported.

BMP7 off-label (cervical)
No comparative studies (as of 8/31/11)

Appendix F. DATA ABSTRACTION TABLES

Appendix Table 1. Comparative studies reported in the AHRQ HTA evaluating BMPs in spinal fusion: study characteristics.

Note. Abstraction tables copied directly from the AHRQ HTA report except that the references and quality of evidence gradings were changed to correspond to the current report. In addition, the applicable key question(s) are noted.

Investigator (yr, country, ref #) Surgical Site Key question	Study design	Comparison(s) No. pts (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of F/U (rng)	Withdrawal or loss to F/U (%)	LoE	Comment
On-label use									
Boden et al., 2000 USA **Lumbar spine** **KQ2, KQ3**	Multicenter, nonblinded RCT	rhBMP2 n=11 (4.2-8.4 mg/pt) ICBG n=3	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	Inclusion: primary symptomatic single-level anterior lumbar fusion, DDD, age 18-65 yrs, grade I spondylolisthesis, symptoms unresponsive to minimum 6 mos. nonoperative therapies Exclusion: spinal condition other than DDD, use of drugs that inhibit bone healing, osteopenia, BMI > 40%, tobacco use, endocrine bone	Radiographic fusion using plain film radiographs and CT analysis, SF-36, Oswestry Low Back Pain Disability Index, neurological functional status, pain medication use, perioperative data, second surgeries, work status, complications and adverse events	24 mos.	0	IIb	Pilot study using rhBMP2 soaked absorbable collagen sponges (ACS) as carrier inside tapered lumbar interbody fusion cages

Investigator (yr, country, ref #) Surgical Site Key question	Study design	Comparison(s) No. pts (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of F/U (rng)	Withdrawal or loss to F/U (%)	LoE	Comment
On-label use									
				disorder					
Burkus et al., 2002 USA **Lumbar spine** **KQ2, KQ3**	Multicenter, nonblinded RCT	rhBMP2 n=143 (4.2-8.4 mg/pt) ICBG n=136	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	Inclusion: primary symptomatic single-level anterior lumbar fusion, DDD, symptoms unresponsive to minimum 6 mos. nonoperative therapies Exclusion: NR	Radiographic fusion using plain film radiographs and CT analysis, Oswestry Low Back Pain Disability Index, neurologic functional status, back, leg and graft site pain numerical rating scales, perioperative data, second surgeries, return to work, complications and adverse events	24 mos.	rhBMP2 20 (14%) ICBG 27 (20%)	IIb	Pivotal trial using rhBMP2 soaked absorbable collagen sponges (ACS) as carrier inside tapered lumbar interbody fusion cages
Burkus et al., 2003 (Integrated analysis)	Retro-spective combined comparative	rhBMP2 n=277 (dose NR)	single-level primary anterior lumbar fusion	Same as Burkus et al., 2002 (72)	Radiographic fusion using plain film radiographs	24 mos.	rhBMP2 30 (11%)	II	Analysis of combined data from 2 published studies (Burkus et

Investigator (yr, country, ref #) Surgical Site Key question	Study design	Comparison(s) No. pts (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of F/U (rng)	Withdrawal or loss to F/U (%)	LoE	Comment
On-label use									
USA **Lumbar spine** Note: may include pts in Burkus et al., 2003[59] ("Radio-graphic assessment …") **KQ2, KQ3**	analysis	ICBG n=402	with interbody fusion cages		and CT analysis, SF-36, Oswestry Low Back Pain Disability Index, perioperative data, second surgeries, work status, complications and adverse events		ICBG 75 (19%)		al., 2002, [72], and Kleeman et al., 2001, [183]) plus unpublished data from a third study. rhBMP2 soaked absorbable collagen sponges (ACS)

Investigator (yr, country, ref #) Surgical Site Key question	Study design	Comparison(s) No. pts (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of F/U (rng)	Withdrawal or loss to F/U (%)	LoE	Comment
Off-label use									
Boden et al., (2002) USA **Lumbar Spine** **KQ2, KQ3**	Multicenter nonblinded RCT	rhBMP2/CRM plus Texas Scottish Rite Hospital (TSRH) Spinal System (TSRHSS) n=11 (40 mg/pt) rhBMP2/CRM alone n=11 (40 mg/pt) ICBG plus TSRHSS n=5	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 ICBG	Inclusion: primary symptomatic single-level lumbar DDD, low back or leg pain unresponsive to minimum 6 mos. nonoperative therapies, grade I or less spondylolisthesis, 18 years or older, Oswestry DI score at least 30 Exclusion: prior fusion at index level, medications that interfere with fusion, scan-confirmed osteoporosis, autoimmune disease, prior exposure to BMP, endocrine	Radiographic fusion using plain film radiographs and CT analysis, Oswestry Low Back Pain Disability Index, SF-36 physical component subscale, neurological functional status, back, leg and graft site pain numerical rating scales, perioperative data, second surgeries, complications and adverse events	mean 17 mos (12-27 mos.)	rhBMP2/CRM alone 2 (18%) were found to have > grade I spondylolisthesis and were excluded from analysis	IIb	IDE pilot study for device which has not received FDA marketing approval Pilot study of rhBMP2 plus an osteoconductive compression-resistant matrix (CRM) composed of 60% hydroxyapatite and 40% tricalcium phosphate bulking agent, plus absorbable collagen sponge (ACS)

				disorders that affect osteogenesis, tumor, infection					
Burkus et al., (2005) USA **Lumbar Spine** Note: includes all pts from Burkus et al., 2002, rec# 11510; same pts as Burkus et al., 2006, rec# 6640 **KQ2, KQ3**	Multicenter, nonblinded RCT	rhBMP2 n=79 (8-12 mg/pt) ICBG N=52	primary single-level anterior lumbar fusion with a pair of threaded allograft cortical bone dowels (CBD) plus rhBMP2 or ICBG	Inclusion: radiographic documentation of primary symptomatic single-level lumbar DDD, age ≥ 18 years, spondylolisthesis grade ≤ 1, symptoms related to neuroradiographic findings unresponsive to minimum 6 mos. nonoperative therapies Exclusion: spinal conditions other than DDD, DDD at disc space levels other than L4-L5 or L5-S-1, previous anterior fusion at index level, obesity (> 40% above ideal wt), active bacterial infection, medication(s) that	Radiographic fusion based on plain film radiographs with use of anteroposterior , lateral, and flexion-extension views, 1-mm slice CT scans with coronal and sagittal reconstructions , Oswestry Low Back Pain Disability Index, SF-36 physical component subscale, back, leg and graft site pain numerical rating scales, work status perioperative data, second surgeries, complications and adverse events	24 mos	rhBMP2 3 (3.8%) ICBG 2 (3.8%)	IIa	rhBMP2 soaked absorbable collagen sponges (ACS)

				could interfere with fusion (e.g., steroids, NSAIDs)					
Dawson et al., 2009 USA **Lumbar spine** **KQ2, KQ3**	Multicenter nonblinded RCT	rhBMP2/CRM n=25 (12 mg/pt) ICBG n=21	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	Inclusion: primary symptomatic single-level lumbar DDD, low back pain or radicular leg pain unresponsive to minimum 6 mos. nonoperative therapies, grade I or less spondylolisthesis Exclusion: NR	Radiographic fusion using plain film radiographs and CT analysis, Oswestry Low Back Pain Disability Index, SF-36 physical component and physical function subscales, neurological functional status, back, leg and graft site pain numerical rating scales, perioperative data, second surgeries, work status, complications and adverse events Overall success defined as	24 mos.	rhBMP2 3 (12%) 1 death, 2 second-surgery failures ICBG 3 (14%) 1 pt without 24 mos. visit, 2 second-surgery failures	IIb	Pilot study for Infuse/Mastergraft device,which has received FDA marketing approval Infuse/Mastergraft comprises rhBMP2, an osteoconductive, compression-resistant matrix (CRM) composed of 15% hydroxyapatite and 85% tricalcium phosphate ceramic bulking agent, plus absorbable collagen sponge (ACS)

					combination of successful fusion, improvement in ODI score > 15%, absence of severe device-related adverse events, no second surgical procedure involving the index level, maintenance or improvement of neurological status				
Dimar et al., (2009) USA **Lumbar Spine** "Note [AHRQ]: contains pts in Glassman et al., 2007,	Multicenter nonblinded RCT	rhBMP2/CRM n=239 (40 mg/pt) ICBG n=224	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	Inclusion: primary symptomatic single-level lumbar DDD, low back pain or radicular leg pain unresponsive to minimum 6 mos. nonoperative therapies, grade I or less	Radiographic fusion using plain film radiographs and CT analysis, Oswestry Low Back Pain Disability Index, SF-36 physical component	24 mos	rhBMP2/CRM 23 (9.6%) ICBG 30 (13%)	IIb	IDE trial for AMPLIFY device, which has not received FDA marketing approval AMPLIFY comprises rhBMP2, an osteoconductive, compression-

rec# 4040; Dimar et al., 2006 rec# 5480; Glassman et al., 2005, rec# 8040" **KQ2, KQ3**				spondylolisthesis, 18 years or older, Oswestry DI score at least 30 Exclusion: prior fusion at index level, medications that interfere with fusion, scan-confirmed osteoporosis, autoimmune disease, prior exposure to BMP or collagen, endocrine disorders that affect osteogenesis, tumor, infection, pregnancy, or inability to harvest bone graft	subscale, neurological functional status, back, leg and graft site pain numerical rating scales, perioperative data, second surgeries, complications and adverse events				resistant matrix (CRM) composed of 15% hydroxyapatite and 85% tricalcium phosphate ceramic bulking agent plus absorbable collagen sponge (ACS)
Glassman et al., (2008) USA **Lumbar Spine** **KQ2, KQ3**	Multicenter nonblinded RCT	rhBMP2 n=50 (dose not reported) ICBG n=52	single- or multi-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	Inclusion: patients > 60 years, primary symptomatic lumbar DDD with spinal stenosis, spondylolisthesis, instability, adjacent level degeneration Exclusion:	Radiographic fusion based on 1-mm slice CT scans with coronal and sagittal reconstructions, Oswestry Low Back Pain DI, SF-36 physical component	24 mos	106 enrolled, 100 (94%) available for 24 mos. F/U 4 excluded (2 from each arm) in perioperative period due to improper	IIb	All patients > 60 years old, but includes those with single- and multi-level DDD, with fusion performed according to each surgeon's preferences using the same

				Not reported	subscale, back and leg pain numerical rating scales		fusion level (1), fusion not performed (1), refusal to follow-up (1), cross-over (1), 2 died		instrumentation rhBMP2 soaked absorbable collagen sponges (ACS) Enrollment not strictly limited to Medicare population
Haid et al., (2004) USA **Lumbar Spine** **KQ2, KQ3**	Multicenter, nonblinded RCT	rhBMP2 n=34 (4.2-8.4) ICBG N=33	single-level primary posterior lumbar interbody fusion (PLIF) with interbody fusion cages plus rhBMP2 or ICBG	Inclusion: symptomatic, single-level lumbar DDD, grade I spondylolisthesis, with disabling low back or leg pain, unresponsive to minimum 6 mos. nonoperative therapies Exclusion: NR	Radiographic fusion based on plain film radiographs with lateral and flexion-extension views, and 1-mm slice CT scans, Oswestry Low Back Pain Disability Index, back, leg and graft site pain numerical rating scales, SF-36 physical component subscale, neurological status, work status perioperative data, second	24 mos	rhBMP2 4 (12%) ICBG 0	IIb	Trial was halted after preliminary CT scans showed bone growth posterior to the PLIF cages, and was not restarted

					surgeries, complications and adverse events				
Glassman et al., (2007) USA **Lumbar Spine** **KQ2, KQ3**	Retro-spective with historical control group	rhBMP2 n=91 (12 mg/pt) ICBG n=35	single- or multi-level primary or revision instrumented posterolateral lumbar fusion	Inclusion: not explicitly delineated Exclusion: not explicitly delineated	Radiographic fusion based on plain film radiographs and 1-mm slice CT scans with coronal and sagittal reconstructions	mn 27 mos (24-38)	f/u NR 91 patients received rhBMP2, only 48 (53%) comparable to ICBG historical controls	III	ICBG historical control group taken from Glassman et al., 2005 (rec# 8040) rhBMP2 soaked absorbable collagen sponges (ACS)
Mumma-neni et al., 2004 USA **Lumbar Spine** **KQ2, KQ3**	Retro-spective single-center cohort study	rhBMP2/AGB n=25 (8.4 mg/pt) ICBG N=19	single- or multi-level primary transforaminal lumbar interbody fusion (TLIF) with interbody fusion cages with rhBMP2 plus AGB or ICBG alone	Inclusion: symptomatic, single-level lumbar DDD, grade I spondylolisthesis, with disabling low back or leg pain, unresponsive to minimum 6 mos. nonoperative therapies Exclusion: NR	Radiographic fusion based on static and dynamic plain film radiographs, modified Prolo Scale that evaluates pain, functional status, economic status, and medication use (Salehi et al., 2004)	mn 9 mos (3-18 mos)	4 of 44 (9)	III	Study compared rhBMP2 in conjunction with ICBG or local autograft bone and ICBG alone
Pradhan et al., 2006 USA **Lumbar Spine**	Prospective consecutive patient single-center cohort study	rhBMP2 n=9 (dose NR) ICBG n=27	single-level primary anterior lumbar interbody fusion (ALIF) with femoral	Inclusion: primary single-level ALIF, low back pain with or without referred leg pain and sciatica,	Radiographic fusion based on plain film radiographs and 1-mm slice CT scans	rhBMP2 mn 26 (rng 23-29) ICBG mn 36	0	III	Reported radiographic and adverse outcomes rhBMP2 soaked absorbable

KQ2, KQ3			ring allograft (FRA) plus rhBMP2 or ICBG	symptoms unresponsive to minimum 6 mos. nonoperative therapies Exclusion: any prior anterior lumbar spine surgery or posterior destabilizing surgery, osteopenia, osteoporosis, osteomalacia, bone growth stimulation		(rng 29-55)			collagen sponges (ACS)
Singh et al., 2006 USA **Lumbar Spine** **KQ2, KQ3**	Prospective single-center case-matched cohort study	rhBMP2/ICBG n=39 (12-36 mg/pt) ICBG N=11	single- or multi-level primary instrumented posterolateral lumbar fusion with rhBMP2 plus ICBG or ICBG alone	Inclusion: radiographic evidence of DDD, grade I-II spondylolisthesis, lower extremity radiculopathy in a defined dermatomal distribution, unresponsive to minimum 6 mos. nonoperative therapies Exclusion: active smokers, prior fusion at the index level(s) malignancy, metabolic bone	Radiographic fusion based on 2-mm slice CT scans with sagittal and coronal reconstructions	24 mos	2 (4.9) from rhBMP2/ICBG group	III	Study compared rhBMP2 in conjunction with ICBG or local autograft bone and ICBG alone Provided radiographic outcomes only

				disease that would preclude instrumentation or inhibit osteogenesis (i.e., Paget disease, osteomalacia, osteogenesis imperfecta), local or systemic bacterial infection, temperature > 38 degrees at surgery, alcohol or drug abuse in treatment, historyof titanium alloy allergy					
Slosar et al., 2007 USA **Lumbar Spine** **KQ2, KQ3**	Prospective consecutive patient single-center cohort study	rhBMP2 n=45 (3-9 mg/pt) ALG N=30	single- or multi-level primary instrumented anterior lumbar interbody fusion (ALIF) with femoral ring allograft (FRA) plus rhBMP2 or allograft bone chips (ALG)	Inclusion: primary single- or multi-level symptomatic DDD, grade I-II spondylolisthesis, unresponsive to minimum 6 mos. nonoperative therapies Exclusion: DDD at > 3 levels, grade > 2 spondylolisthesis, tumor, infection, psychological contraindications	Radiographic fusion based on plain film radiographs and CT scans, Oswestry Low Back Pain Disability Index, Numerical Rating Scale (NRS) for pain (location not specified)	24 mos	rhBMP2 2 (4) ALG 1 (3)	III	FRA inserts used instead of interbody fusion cages to contain rhBMP2 on ACS or ALG

Johnsson et al., 2002 Sweden **Lumbar Spine** **KQ2, KQ3**	Multicenter nonblinded RCT	rhBMP7 n=10 (7 mg/pt) ICBG n=10	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	Inclusion: radiographic evidence of lumbar DDD, L5 spondylolisthesis, maximal vertebral slip of 50%, intractable lumbosacral pain unresponsive to 6 mos. nonoperative therapies, no radiating leg pain, age > 20 years Exclusion: NR	Radiographic fusion with plain film radiographs, radiostereometric analysis (RSA), patient's subjective evaluation of back pain	12 mos	0 lost to f/u 1 (declined to enroll)	IIb	Efficacy study compared rhBMP7 (OP-1 Putty) and ICBG, based on RSA results
Kanayama et al., 2006 Japan, USA **Lumbar Spine** **KQ2, KQ3**	Multicenter nonblinded RCT	rhBMP7 n=9 (7 mg/pt)	single-level primary instrumented posterolateral lumbar fusion with rhBMP7 or AGB/CRM	Inclusion: radiographic evidence of lumbar DDD, grade I spondylolisthesis with stenosis, neurogenic claudication, unresponsive to minimum 3 mos. nonoperative therapies, age < 85 years	Radiographic fusion with plain film radiographs and CT scan, surgical exploration of fusion mass, Oswestry Low Back Pain DI	rhBMP7 mn 16 mos	rhBMP7 1 (declined to complete study)	IIb	rhBMP7 Putty (OP-1 Putty) compared to local autograft bone admixed with hydroxyapatite plus tricalcium phosphate biphasic cerami cgranules

		AGB/CRM n=10		Exclusion: > 5 degrees kyphosis in flexion, history of fusion at index level, active spinal or systemic infection, known sensitivity to any component of the BMP device, pregnancy or lactation, possible need for additional lumbar surgery within 6 mos		AGB mn 13 mos			
Vaccaro, Lawrence, et al., 2008 USA **Lumbar Spine** **KQ2, KQ3**	Multicenter, nonblinded RCT	rhBMP7 n=207 (7 mg/pt)	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	Inclusion: radiographic evidence of lumbar DDD grade I or II lumbar spondylolisthesis, neurogenic claudication, unresponsive to minimum 6 mos. nonoperative therapies, skeletally mature Exclusion: > Grade II spondylolisthesis, nondegenerative spondylolisthesis of any grade, spinal instability on flexion-extension	Primary Overall Success at 24 mos, a composite measure that required success in all of the following: a 20% improvement in Oswestry Low Back Pain DI, absence of treatment-emergent serious adverse events related to the device, absence of a decrease in neurologic	rhBMP7 mn 53 mos (44-65)	335 enrolled and randomized, 295 (88%) were treated rhBMP7 20 voluntarily withdrew or were disqualified based on the inclusion and exclusion criteria	IIb	IDE study for rhBMP7 device (OP-1 Putty) that did not receive FDA marketing approval Summarize data from 36+ mos. F/U
		ICBG n=86				ICBG 54 (45-66)	ICBG 20 refused autograft or did not qualify after		

				radiographs with > 50% translation of vertebral body or > 20 degrees of angular motion, active spinal or systemic infection, systemic disease precluding participation (eg, neuropathy), current nicotine use, history of smoking, morbid obesity, known sensitivity to collagen	status (assessing muscle strength, reflexes, sensation, and straight leg raise), and radiographic fusion success Modified Overall Success at 36 + mos, a composite measure that required success in all of the following: a 20% improvement in Oswestry Low Back Pain DI, absence of treatment-emergent serious adverse events related to the device, absence of a decrease in neurologic status (assessing muscle		randomization based on the inclusion and exclusion criteria		

					strength, reflexes, sensation, and straight leg raise) at 24 mos, and radiographic fusion success indicated by CT evidence for the presence of new bone, angulation ≤ 5 degrees, translation movement ≤ 3 mm on flexion/extension radiographs, and absence of retreatment to promote fusion at 36+ mos				
Vaccaro et al., 2008 USA **Lumbar Spine** Note: Long-term F/U study that includes all pts from Vaccaro et al., 2004,	Multicenter, nonblinded RCT	rhBMP7 n=24 (7 mg/pt) ICBG n=12	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	Inclusion: radiographic evidence of lumbar DDD grade I or II lumbar spondylolisthesis, neurogenic claudication, unresponsive to minimum 6 mos. nonoperative therapies, minimum Oswestry Low Back Pain	Radiographic fusion based on anteroposterior, lateral, and dynamic flexion-extension lateral plain film radiographs Oswestry Low Back Pain DI, SF-36 physical	48 mos	Radiographic results rhBMP7 9 (38%) Clinical results rhBMP7 5 (21%) Radiographic results ICBG 6 (50%)	IIb	IDE study for rhBMP7 device (OP-1 Putty) that did not receive FDA marketing approval

and Vaccaro et al., 2005 **KQ2, KQ3**				Disability Index score 30 Exclusion: prior lumbar fusion or ICBG harvesting, active infection, history of tobacco use, morbid obesity, known sensitivity to collagen, grade III or IV spondylolisthesis, > 20% angular motion of the listhetic segment	and mental componemt subscales, adverse events and complications		Clinical results ICBG 5 (42%)		
Baskin et al., 2003 USA **Cervical Spine** **KQ2, KQ3**	Multicenter, nonblinded RCT	rhBMP2/ALG n=18 (0.6-1.2 mg/pt) ICBG/ALG n=15	single- or two-level primary instrumented ACDF with rhBMP2/ALG or ICBG/ALG	Inclusion: primary symptomatic single- or two-level cervical DDD with radiculopathy, myelopathy, or both, herniated disc, posterior osteophytes or both at index level(s), symptoms unresponsive to minimum 6 mos. nonoperative therapies Exclusion: NR	Radiographic fusion using plain film radiographs and CT analysis, Neck Disability Index, neck and arm pain, SF-36 physical and mental component subscales, neurologic status (motor and sensory function), patient satisfaction, complications	24 mos	Radio-graphic: 13 (39%) Clinical: 10 (28%)	IIb	Pilot study using rhBMP2 soaked ACS packed inside fibular allograft (ALG) bone **Follow-up data corrected by Spectrum**

					and adverse events				
Butterman et al., 2008 USA **Cervical Spine** **KQ2, KQ3**	Prospective nonrandomized cohorts of consecutive patients	rhBMP2/CRA n=30 (0.9-3.7 mg/pt) ICBG n=36	single- or multi-level primary instrumented or uninstrumented ACDF with rhBMP2/CRA or ICBG	Inclusion: primary symptomatic single- or multi-level cervical DDD Exclusion: Prior ACDF at any level, corpectomy, deformity, presence of tumor, inflammatory joint disease, or cervical spine discitis	Radiographic fusion using plain film radiographs and high-resolution CT, Oswestry Neck Disability Index, neck and arm pain, pain medication use, patients' overall opinion of treatment success	24-36 mos	0	III	rhBMP2/ACS was placed inside the CRA, with resected osteophytes and local bone shavings, compared to ICBG alone
Crawford et al., 2009 USA **Cervical Spine** **KQ2, KQ3**	Retrospective cohort of consecutive patients	rhBMP2/BGE n=41 (4.2-12 mg/pt) ICBG n=36	single- or multi-level instrumented posterior cervical spinal fusion with rhBMP2/BGE or ICBG	Inclusion: single- or multi-level symptomatic posterior cervical stenosis, ACDF non-union, or segmentally unstable spondylosis Exclusion: acute trauma, infection, presence of tumor, concomitant anterior fusion	Perioperative complications, surgical data	≤ 3 mos	0	III	rhBMP2/ACS was combined with bone graft extenders (BGE) including local autograft bone, allograft, or ceramics
Smucker et al., 2006[55] USA	Retrospective case-control	rhBMP2/CRA n=69 (dose NR)	single- or multi-level instrumented	Inclusion: NR	Cervical swelling complications	≤ 6 wks	NR	III	Most patients received cortical ring allograft

Cervical Spine **KQ2, KQ3**		CRA n=165	ACDF with rhBMP2/CRA or CRA alone	Exclusion: NR					(CRA) (88% with rhBMP, 81% of controls)
Vaidya, Carp, et al., 2007 USA **Cervical Spine** **KQ2, KQ3**	Retro-spective cohorts of consecutive patients	rhBMP2 n=22 (1-3 mg/pt) ALG/DBM n=24	single- or multi-level primary instrumented ACDF with interbody fusion cages rhBMP2 on ACS or ALG/DBM	Inclusion: primary symptomatic single- or multi-level cervical DDD amenable to ACDF Exclusion: Prior ACDF at index level(s), trauma, presence of tumor, those more amenable to posterior surgery or combined surgery	Radiographic fusion using plain film radiographs and CT, Oswestry Neck Disability Index, arm and neck pain, perioperative outcomes and complications including swelling, hoarseness, and dysphagia	24 mos	12 (21%)	III	rhBMP2/ACS was placed in polyetheretherketone (PEEK) interbody fusion cages, compared to use of allograft (ALG) spacers with demineralized bone matrix (DBM)

Appendix Table 2. Comparative studies reported after the AHRQ HTA search period evaluating BMPs in spinal fusion: study characteristics.

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
On-label use									
FDA SSED: InFUSE (P000058) **Lumbar spine** **(**overlaps with Boden 2000 RCT, Burkus 2002 RCT, Burkus 2003 integrated analysis) **KQ3**	Integrated analysis (of pilot (Boden 2000[13]) and pivotal (Burkus 2002[14] + subset of Burkus 2003[15]))	rhBMP-2: n = 288 ICBG: n = 139	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	Inclusion: DDD with back pain with or without leg pain at a single level between L4 and S1 confirmed by history and radiographic studies. DDD present if one or more of the following were noted: instability, osteophyte formation, decreased disc height, ligament thickening, disc degeneration/ herniation, or facet joint degeneration. In addition, the following were required: pre-op ODI score of 35+, spondylolisthesis grade 1 (if present) non-responsive to non-operativetreatment for at least 6 months, skeletally mature, and not pregnant or nursing and agrees to the use of contraception for 16+ weeks post-implantation. Exclusion:	Adverse events	< 30 months (range, mean f/u NR)	NR	n/a	**Study funding:** Both the pilot and pivotal trials were sponsored by the manufacturer of InFUSE (Medtronic)

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
On-label use									
				Previous anterior spinal fusion at the involved level, posterior spinal instrumentation at the involved level or a previous interbody fusion procedure, any conditions that require postop medications that would be expected to interfere with fusion, osteopororsis, osteopenia, or osteomalacia, active malignancy, active local or systemic infection, gross obesity (>40% ideal body weight), fever > 101°F, mentally incompetent, Waddell Signs of Inorganic Behavior ≥ 3, alcohol or drug abuse, tobacco user, autoimmune disease, titamium allergy, previous exposure to injectable collagen implants, hypersensitivity to protein pharmaceuticals or collagen, previous exposure to rhBMP-2, allergy to bovine					

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
On-label use									
				products or history of anaphylaxis, endocrine or metabolic disorder that affects osteogenesis, or received another investigational therapy within 28 days prior to implantation.					

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
Burkus et al. (2011) USA **Lumbar spine** **KQ3** (patients from FDA SSED Pivotal Study; also reported in Burkus 2002 and subset of Burkus 2003 integrated analysis, Dimar 2009 RCT, as well as from Gornet 2007 RCT (abstract only))	Cohort study: integrated analysis of 3 studies	BMP-2: n = 1093 (varying surgical interventions) (dose NR) Autograft (ICBG): n = 360	Study #1 (on-label use) (patients from FDA SSED Pivotal Study; also reported in Burkus 2002 and subset of Burkus 2003 integrated analysis): ALIF with LT-CAGE done laproscopically (n = 134, BMP only, nonrandomized arm) or with open surgery (BMP2, n = 143; ICBG, n = 136, randomized arm) Study #2 (Gornet 2007 RCT): open ALIF with BMP (all pts) using lumbar tapered fusion device (n = 172) (on-label use) or metal-on-metal lumbar disc arthroplaty device (n =	Inclusion: Single-level symptomatic DDD, grade I spondylolisthesis or lower, or disabling back and/or leg pain unresolved by nonoperative treatment for longer than 6 mos. Women of childbearing age asked to delay any pregnancies following surgery by 16 weeks- 12 months. Exclusion: Spinal conditions other than DDD, previous anterior or posterior fusion at the involved level, obse (>40% above ideal body weigh), active bacterial infection, medical condition requiring medication that might interfere with fusion.	Antibody responses, correlation with fusion, adverse events, and miscarriages	Varied: Study #1: 3 mos. Studies #2 & 3: 1.5, 3, 6, 12 mos.	NR	III	**A positive antibody response is present when:** (1) the baseline sample is negative and any post-treatment sample is positive (titer ≥ 50); (2) the baseline sample is positive and any post-treatment samples have titers 2–3X higher than the baseline titer (depending on the assay used); or (3) the baseline sample is unavailable and any post-treatment sample is positive. **Study funding:** Medtronic Spinal & Biologics

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
			405) (off-label use). Study #3 (off-label use) (Dimar 2009): single-level instrumented posterolateral lumbar arthrodesis through open approach with BMP-2-matrix (n = 239) or ICBG (n = 224).						
Carragee et al. (2011) USA **Lumbar spine** **KQ3**	**Retrospective cohort**	rhBMP2: n = 69 (4.2 mg/pt) Osteophytes or ICBG: n = 174	1- or 2-level ALIF including L5/S1 via an open retroperitoneal approach with a FRA or titanium mesh cage filled with ICBG or	Inclusion: 1- or 2- level ALIF; degenerative spondylolisthesis, low-grade isthmic spondylolisthesis, recurrent lumbar disc herniation, or presumed discogenic pain; lumbar fusion	Radiographic: none Clinical: Retrograde ejaculation	12 mos. (early posop, 12 mos.)	0% (0/243)	III	If rhBMP-2 was used, two sponges (4.2 mg) were placed inside the FRA central canal; unless a four-hole plate was used in a stand-alone configuration, a

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
			rhBMP-2/ACS; posterior instrumentation used at discretion of surgeon.	crossed 1 or 2 disc levels and included the L5/S1 level Exclusion NR					buttress screw was placed (into the caudal vertebrae just below the end plate) **Study funding:** No funds received or will be received
Crawford et al. (2010) USA **Sacrum** **KQ2, KQ3** (appears to contain the same patients reported in Maeda (2009))	Retrospective cohort with historical control	rhBMP2: n = 39 (dose NR) Autogenous graft (iliac crest, rib, or local) (historical controls): N = 25	Posterior extension of an existing fusion to the sacrum with segmental pedicle screw instrumentation, including S1 pedicle screw fixation and iliac screw fixation; all but five patients (study group) had anterior interbody device support at the lowest level via an anterior or transforaminal approach	Inclusion: Patients who had undergone long idiopathic scoliosis fusion as an adolescent or young adult and later presented with distal degeneration requiring extension of the fusion to the sacrum Exclusion: NR	Radiographic: coronal and sagittal imbalance; thoracic Cobb angle; lumbar Cobb angle; T5–T12 sagittal Cobb angle; T10–L2 sagittal Cobb angle; T12-sacrum sagittal Cobb angle; segmental lordosis (Cobb angle) from end of previous fusion to sacrum; fusion/ nonfusion; pseudarthrosis Clinical: SRS-22 preoperative, SRS-30	≥ 2 years rhBMP2: 3.3 ± 2.2 years Autogenous graft: 5.1 ± 1.9 years	rhBMP: 7.7% (n = 3); (92.3% follow-up) Autogenous graft: 4.0% (n = 1) (94% follow-up)	III	Fusions were evaluated by two independent spine surgeons; no other mention of independent assessment **Study funding:** No funds received to support the study; however one or more authors has/have received or will receive(d) funds from commercial parties related to the study

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
					postoperative, ODI, medical and surgical complications				
Howard et al. (2011) USA **Lumbar spine** **KQ3**	Cross-sectional	rhBMP2: n = 59 (dose NR) ICBG: n = 53	1- to 2- level instrumented posterolateral fusion from L1 to S1 NOTE. ICBG harvested through midline lumbar incision (no scar over graft site).	Inclusion: 1- to 2- level instrumented posterolateral fusion from L1 to S1 Exclusion: Possible or definite pseudoarthrosis based on imaging studies or fusion extending into the thoracic spine	Clinical: incidence and severity of bone graft site pain	41 months (6–211)	NR	III	Patients assessed by independent investigator, not directly involved in the care of the patient and unaware of the type of bone graft used in the fusion The patients were asked to rate the intensity of the pain with direct palpation over each crest on a scale of 0 to 10 with 0 being no pain and 10 being the worst pain **Study funding:** Funds were received by more than one author from commercial parties related to the study (not clear if there was direct funding of the study) (Medtronic, Norton

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
									Healthcare)
Joseph et al. (2007) [32 32 31 30] USA **Lumbar spine** **KQ3**	Prospective cohort	rhBMP2: n = 23 (24 levels) (4.2 mg/level) Local autograft: n = 10 (12 levels)	Minimal access PLIF or TLIF with interbody cages and percutaneous pedicle screw fixation	Inclusion: Cohort of consecutive patients who had undergone posterior minimal access PLIF or TLIF fusion. BMP use was "ultimately" at patient's discretion. Exclusion: NR	Heterotopic bone formation, other complications	Radiographic: 7.9 (6-16) mos. Clinical: 25.0 (18-52) mos.	3% (1/34)	III	AHRQ considered this a case series, so will be evaluated for KQ3 only. CT scans done prospectively. **Study funding:** Direct funding not received; authors received royalities, consulting fees, speaking arrangements, trips/travel, had stock ownership, and/or were on the scientific advisory board from/of DePuy Spine, Medtronic, Inuve Gertis, and/or Syntheses.
Latzman et al. (2010) USA **Lumbar spine**	Retrospective cohort	rhBMP2: n = 24* (12 mg/8 cc; 24 mg/16 cc) Auto- or allograft only: N = 105*	Lumbar and lumbosacral spinal fusion with and without interbody cage placement	Inclusion: All patients undergoing lumbar or lumbosacral fusion between July 1, 2000 and June 23, 2008. Exclusion: Concordant or prior	Clinical: Renal insufficiency, complications, new diagnoses Radiographic: NR	rhBMP2: mean 1.5 ± 0.85 years Auto- or allograft only: mean 4.5 ± 2.0 years	NR	III	**Study funding:** No direct support but one or more authors has/have received or will receive monetary benefits from commercial party related directly or

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
KQ3				resection of lumbar and lumbosacral tumors					indirectly to manuscript
Lee et al. (2010) USA **Lumbar spine** **KQ2, KQ3**	Retrospective cohort	rhBMP2 (with allograft): n = 86 (4.2 mg/2/8 ml for 1-level; 8.4 mg/5.6 ml for 2-level; 12 mg/8.0 ml for 3+ levels) age ≥ 65 years: n = 34 age < 65 years: n = 52 ICBG: age ≥ 65 years: n = 41	Instrumented posterolateral lumbar fusion (PLF)	Inclusion: Instrumented PLF for the treatment of degenerative lumbar spine diseases between 2002 and 2006; procedure utilized rhBMP-2 with allograft or autograft only; ≥ 2 years of follow-up Exclusion: anterior or posterior lumbar interbody fusion, an uninstrumented PLF; < 2 years of follow-up	Radiographic: Fusion rate, fusion time Clinical: Results according to Kirkaldy-Willis criteria, VAS pain, perioperative complication rates, revision rates	rhBMP2 age ≥ 65 years: 38.3 ± 7.4 mos. (24–68) age < 65 years: 39.2 ± 11.7 mos. (24–62) ICBG age ≥ 65 years: 34.7 ± 8.2 mos. (24–58)	NR	III	**Study funding:** NR
Rihn et al. (2009) USA **Lumbar spine** **KQ3**	Retrospective cohort	rhBMP2 n = 86 ICBG n = 33	TLIF 1-level Primary or revision	Inclusion: Patients 18-80 years of age who underwent single-level TLIF using either rhBMP-2 or ICBG for treatment of a degenerative condition including degenerative or isthmic spondylolisthesis and/or had prior lumbar surgery. Exclusion:	Complications, ICBG donor site pain (assessed using a questionnaire via telephone interview).	mean 24.4 mos. mean 35.8 mos. (*P* < .001)	8.4% (11/130)	III	AHRQ considered this a case series, so will be evaluated for KQ3 only. **Study funding:** Direct funding NR; authors received royalities, consulting fees, speaking arrangements, trips/travel, had stock ownership,

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
				Patients who underwent a multilevel TLIF procedure, who received a bone graft substitute or extender other than rhBMP2, or who had operative treatment for nondegenerative conditions (ie., tumor, infection , or trauma).					and/or were on the scientific advisory board from/of DePuy Spine, Medtronic, Inuve Gertis, and/or Syntheses.
Taghavi et al. (2010) USA **Lumbar spine** **KQ2, KQ3**	Retrospective cohort	rhBMP2 n = 24 (1.5 mg/mL concentration; 12 mg regardless of no. of levels) BMAA N = 18 Autograft N = 20	Transpedicular instrumented revision posterolateral fusion; rhBMP2 (INFUSE kit, 12 mg, 1.5 mg/mL concentration, ACS); BMMA from a single iliac crest; autograft	Inclusion: Instrumented revision posterolateral fusion between January 2002 and December 2006; minimum 2-year follow-up; symptomatic pseudarthrosis following previous posterolateral fusion for DDD Exclusion: Infection, tumor, trauma	Radiographic: Fusion rate, time to solid fusion, nonunion Clinical: VAS for back and leg pain, complications	rhBMP2 28.4 mos. BMMA 27.6 mos. Autograft 27.6 mos.	NR	III	Two spine surgeons blinded to the graft material used and an independent consultant radiologist evaluated the progression of the fusion mass **Study funding:** No direct support but one or more authors has/have received or will receive monetary benefits from commercial party related directly or indirectly to manuscript

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
Vaidya, Weir et al. (2007) USA **Lumbar and cervical spine** **KQ3**	Prospective cohort	rhBMP2 + allograft: n = 36 (55 levels)	rhBMP2 + allograft (in cages): ALIF: n = 13 (20 levels) TLIF: n = 12 (17 levels) Anterior cervical decompression /fusion: n = 11 (18 levels)	Inclusion: Consecutive patients who required a cervical or lumbar interbody fusion. Exclusion: NR	Nonunion, early lucency, subsidence	24.1 (17-30) mos.	0% (0/77)	III	AHRQ excluded this study (as "not relevant design"), so will be evaluated for KQ3 only. **Study funding:** No benefits received or will be received from a commercial party.
		DBM + allograft: n = 41 (63 levels)	DBM + allograft (in cages): ALIF: n = 11 (16 levels) TLIF: n = 18 (25 levels) Anterior cervical decompression /fusion: n = 12 (22 levels)			24 (18.5-27) mos.			
Delawi et al. (2010) Europe (Netherlands, France, Italy, Spain)	**RCT** Multicenter (5)	OP-1 (rhBMP-7): n = 18 (3.5 mg per side of spine)	Primary, 1-level, posterolateral lumbar fusion using pedicle screw instrumentatio n;	Inclusion: Degenerative or isthmic spondylo (grades I and II) with central or foraminal stenosis; Eligible for decompression and single-level fusion	Radiographic: Fusion Clinical: ODI; donor site pain for ICBG group (VAS, 1-10);	12 months (NR)	89% (32/36)	IIb	CT scans were reviewed by a spinal surgeon and a senior radiology resident blinded to the treatment group and the institute

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
Lumbar spine **KQ2, KQ3**		Autograft (ICBG): n = 16	decompression via bilateral laminectomy or partial laminectomy and medial facetectomy; under general anesthesia; prophylactic cephalosporin given for 24 hours starting 15 mins. before incision	(L3–S1); Symptoms of radiculopathy or neurogenic claudication; A preoperative ODI > 30; Nonresponsive to at least 6 months of nonoperative treatment; No previous fusion attempt(s) to the affected level; Skeletally mature Exclusion: Gross instability that requires multiple levels fusion; Severe osteoporosis or osteopenia; Suspicion of active spinal or systemic infections; Women who were pregnant or who planned to become pregnant; Known sensitivity to collagen; Morbid obesity; Patients who have in the last year been prescribed systemic corticosteroids; Known to require additional surgery to the lumbar spinal region within 6 months	safety/adverse events Clinical assessments done at 6 wks. and 3, 6, and 12 months				where the procedure was performed. A third observer, a spinal surgeon, was used to adjudicate conflicting findings. **Study funding:** Corporate/ industry and institutional funds were received in support of the work; specific source(s) NR.

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
Hwang et al. (2010) USA; Canada **Lumbar spine** **KQ3** (same patients as Vaccaro, Lawrence (2008); additional safety data reported)	**RCT** Multicenter (24)	OP-1 (rhBMP-7): Initial phase (f/u to 24 mos): n = 228 (only 208 of whom were treated) Extended phase (for 36+ mo f/u): n = 144 enrolled Autograft: Initial phase (f/u to 24 mos): n = 108 (only 87 of whom were treated) Extented phase (for 36+ mo f/u): n = 58 enrolled	Single-level decompression and uninstrumented posterolateral fusion of the listhetic segment	NR	<u>Radiographic:</u> Fusion <u>Clinical:</u> Anti-OP-1 antibodies (Nabs or neutralizing antibodies); overall success, ODI improvement ≥ 20%, neurological success, absence of retreatment, absence of treatment-emergent serious adverse events	Mean 4.4 years	rhOP-1 6 weeks: 84.5% (284/336) 3 months: 81.5% (274/336)% 6 months: 82.7% (278/336)% 12 months: 77.1% (259/336)% 24 months: 70.5% (237/336) Extended phase of study (36+ months): 67.3%	IIb	Objective of the paper was to examine the presence and effect of OP-1 Nabs on the safety and efficacy of rhOP-1; thus the analysis of the effect on fusion and clinical success included only the patients treated with OP-1 putty. Some safety data reported for OP-1 vs. autograft. **Study funding:** Stryker Biotech
Xu et al. (2011) USA **Cervical spine** **KQ2, KQ3**	Retrospective cohort study	rhBMP-2 + some/all of the following (DBM (31%), local autograft (77%), allograft (21%), hydroxyapatite crystals (61%): n = 48 (dose NR)	Primary posterior cervical arthrodesis (single- or multi-level) (mean 5.9 ± 1.9 levels/pt)	<u>Inclusion:</u> Consecutive patients undergoing primary posterior cervical arthrodesis for symptomatic primary degenerative cervical pathologies. <u>Exclusion:</u> Trauma, tumor,	Intraoperative blood loss, length of stay, fusion, neck pain, Nurick score, ASIA score, adverse events, reoperation	24.2 ± 10.1 months (range, 1-39.6 mos)	17.1% (35/204)	III	**Study funding:** NR

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
		Non-BMP: some/all of the following (DBM (86%), local autograft (88%), allograft (72%), hydroxyapatite crystals (0%): n = 156		infections, fusion of only C1-C2, systemic metabolic disorders that secondarily affect bone quality such as renal osteodystrophy.					
Yaremchuk (2010) USA **Cervical spine** **KQ2, KQ3**	Retrospective cohort study	BMP (n = 260) -dosages NR Non-BMP (n = 515)	Cervical spinal fusion with or without BMP (approach NR)	NR	Length of stay (LOS), hospital charges, incidence of airway obstruction, unplanned intubations after surgery, tracheotomies, ICU admissions, hoarseness, dyspnea, respiratory failure, dysphasia, dysphagia, hospital readmissions, need for percutaneous endoscopic gastrostomy (PEG) tubes, death	30 d for all measures except death (90 d postop identified)	NR	III	Type of BMP used not specified

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
Cahill (2011) USA **Lumbar spine** **KQ2, KQ3**	Retrospective case control study	BMP (rhBMP-2 OR rhBMP-7) ± autograft ± allograft n = 2,372 (6% received autograft harvested from a different incision) (28% received allograft)	Single-level lumbar fusion (any approach) with or without BMP (rhBMP-2 or rhBMP-7). Fusion type: Interbody: 35% Posterolateral: 18% Circumefer-ential: 48% Instrumented fusion: 87% (%s similar in both groups)	Inclusion: Patients (>18 years of age) in the MarketScan database who underwent a single-level lumbar fusion between 2003 and 2008 and had at least one-year follow-up. Patients identified using CPT-4 and corresponding ICD-9 codes for interbody, posterolateral, or circumferential (both an interbody and posterolateral) fusion. Exclusion:	Repeat lumbar fusion, postoperative inpatient length of stay, risk of 30-day repeat inpatient admission.	12 months minimum BMP: mean 2.18 ± 0.98 yrs No BMP: mean 2.19 ± 0.99 yrs	NR	III	**Database:** MarketScan Commercial Claims and Encounters database (Thomson Reuters Inc.): longitudinal health insurance dataset taken from inpatient and outpatient settings and yearly enrollment data. Includes administrative claims from ~100 insurance companies and

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
		No BMP ± allograft (matched controls): n = 2,372 (26% received alloigraft) Propensity scores used to match patients that underwent fusion with BMP to controls with a similar probability of undergoing a fusion with BMP.		Dianoses related to spinal cancer, infectious processes, or trauma.					large employers; represents >69 million patients since 1996. **Patients who met inclusion criteria:** 15,862 pts with one-year follow-up (out of total pool of 21,216 pts); 2373 pts received BMP and 13,489 underwent fusion without BMP. **Medical comorbidity stratification** done using the Charlson comorbidity index; Charlson scores determined by averaging all inpatient admissions during the immediate 3 months prior to and including the index procedure. Other clinical comorbidities (ie., osteoporosis, obesity, diabetes,

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
									tobacco use) identified from inpatient and outpatient records. **Study funding:** Harvard Catalyst/ The Harvard Clinical and Translational Science Center (NIH award); Harvard University and affiliated academic health care centers.
Deyo et al. (2011) USA	**Retrospective cohort (database) study**	BMP: n = 1703	Single- or multilevel, primary or repeat fusion.	Inclusion: Patients ≥ 68 years of age in the MedPAR database who received an index	Complications	≥ 4 yrs (specifics NR)	NR	III	**Database:** Medicare Provider Analysis and Review (MedPAR)

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
Lumbar spine **KQ3**		No BMP: n = 15,119		procedure in 2003 or 2004 for stenosis who were Medicare beneficiaries eligible through the Old Age and Survivors Insurance program. Exclusion: Patients receiving Social Security Disability Income, end-stage renal disease, patients enrolled in health maintenance organization at time of the index visit. Patients with codes indicating cancer, vehicular accident, spinal infection, inflammatory spondylo-arthropathies, vertebral fractures or dislocations, or cervical or thoracic spine procedures.					database, which includes all Medicare hospital claims. Data on mortality taken from another file maintained by the Centers for Medicare and Medicaid Services. **Simple fusion:** anterior fusion, transverse process, OR posterior fusion with 1-2 levels (2-3 vertebrae). **Complex fusion**: 360° fusion by single incision, combination of anterior with either transverse process or posterior fusion techniques, or any fusion with 3 or more levels.

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
Mines et al. (2011) USA **Lumbar spine** **KQ3**	Retrospective cohort (database) study	rhBMP-2: n = 15,640	Lumbar fusion surgery with or without BMP2	Inclusion: Medicare patients (≥67 years of age) who underwent lumbar fusion surgery between Oct 2003 and Dec. 2005 who were continuously enrolled in fee-for-service Medicare for at least 2 years prior to the index procedure. Exclusion: Claim for pancreatic cancer within 2 years prior to index procedure; participants in Medicare- funded HMOs, patients without continuous participation in Medicare Part B, patients covered by Medicare due to end-stage renal disease or chronic disability.	Pancreatic cancer	BMP: median of 0.91 (IQR, 0.41, 1.54) years (45.56% had ≥12 mos. f/u)	BMP: 3.1% (deaths)	III	**Database:** Medicare claims data from 3 sources: Medicare Provider Analysis and Review (MEDPAR) file (includes services provided in Medicare-certified inpatient hospitals); Carrier file (claims from physicans and free-standing ambulatory surgical centers); and Outpatient file (includes claims from outpatient providers, including outpatient hospital visits).
		No BMP: n = 78,194				No BMP: median of 1.47 (IQR, 0.73, 2.21) years (65.76 % had ≥12 mos. f/u)	No BMP: 5.1% (deaths)		
Cahill et al. (2009) USA **Lumbar, cervical, or thoracic spine**	Retrospective cohort (database) study	BMP (any): n = 17,623	Fusion (any) BMP: Revision fusion: 8.52% (1502/17,623) Cervical: 16.38% (2886/17,623)	Inclusion: Patients (> 18 years of age) in the Nationwide Implant Sample database who underwent a primary or revision fusion in 2006. Patients identified using ICD-9 codes	Complications	Duration of inpatient stay	NR	III	AHRQ excluded this as a cost study only, so will be evaluated for KQ3 only. **Database:** Nationwide Implant Sample database

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
KQ3			Lumbosacral: 79.28% (13,972/ 17,623) Thoraco-lumbar: 4.23% (746/17,623) Unknown: 0.11% (19/17,623) Vertebral levels, 2-3: 83.03% (14,633/ 17,623) Vertebral levels, ≥4: 16.97% (2990/17,623)	primary and revision fusions fusion. <u>Exclusion:</u> NR					(nationwide sample of hospital discharge records) (part of the Healthcare Cost and Utilization Project), contains data from 5–8 million discharges per year from sample of hospitals (~20% of US hospitals), and includes all payers. **Medical comorbidity stratification** done using the Charlson comorbidity index. **Funding:** Brain Science Foundation, which had no role in the design and conduct of any part of the study. No financial disclosures reported
		No BMP: n = 53,026	<u>No BMP:</u> Revision fusion: 4.89% (2595/53,026) Cervical: 52.03% (27,589/ 53,026) Lumbosacral:						

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Surgical intervention	Inclusion/exclusion criteria	Outcomes measured	Duration of f/u (range)	Withdrawal or loss to f/u (%)	LoE	Comment Study funding or sponsorship
Off-label use									
			43.06% (22,835/53,026) Thoracolumbar 4.74% (2511/53,026) Unknown: 0.17% (91/53,026) Vertebral levels, 2-3: 84.57% (44,846/53,026) Vertebral levels, ≥4: 15.43% (8180/53,026)						

ACS: absorbable collagen sponge; BMAA: bone marrow aspirate with allograft; DDD: degenerative disc disease; DBM: demineralized bone matrix; FDA: Food and Drug Adminstration; f/u: follow-up; FRA: femoral ring allograft; ICBG: iliac crest bone graft; IDE: investigational device exemption; LoE: level of evidence; ODI: Oswestry Low Back Pain Disability Questionnaire; PLF: posterolateral lumbar fusion; PLIF: posterior interbody fusion; RCT: randomized controlled trial; rhBMP2: recombinant human bone morphogenetic protein 2; SF-36: Short-Form 36; TLIF: transforaminal lumbar interbody fusion

*Lumbar and lumbosacral fusion was performed on 125 patients; 101 patients underwent 104 operations without rhBMP2 and 20 underwent 23 operations with rhBMP2. Four patients had 1 operation with rhBMP2 and 1 without rhBMP2, for a total of 8 operations. There were 135 total operations.

Appendix Table 3. Comparative studies reported in the AHRQ HTA evaluating BMPs in spinal fusion: patient demographics.

Note. Abstraction tables copied directly from the AHRQ HTA report except that the references and quality of evidence gradings were changed to correspond to the current report.

Investigator (yr, country, ref #) Surgical Site	Study design	Comparison(s) No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD yrs (rng)	≥ 65 yrs (%)	Males (%)	Weight mean ± SD lbs (rng)	Comorbidities (%)	Comment
On-label use											
Boden et al., 2000 USA **Lumbar spine** **KQ2, KQ3**	Multicenter, nonblinded RCT	rhBMP2 (4.2-8.4 mg/pt) n=11 ICBG n=3	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	grade I spondylolisthesis	rhBMP2 42±3 (30-62) ICBG 40±0.6 (38-42)	NR	rhBMP2 46 ICBG 67	rhBMP2 166±11 (125-228) ICBG 211±11 (190-249)	Tobacco use rhBMP2 0 Frequent alcohol use rhBMP2 36.4 Tobacco use ICBG 33.3 Frequent alcohol use ICBG 33.3	No significant differences between groups
Burkus et al., 2002 USA **Lumbar spine** **KQ2, KQ3**	Multicenter, nonblinded RCT	rhBMP2 (4.2-8.4 mg/pt) n=143 ICBG n=136	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	NR	rhBMP2 43 ICBG 42	NR	rhBMP2 54 ICBG 50	rhBMP2 179 ICBG 181	Tobacco use rhBMP2 33 ICBG 36	No significant differences between groups
Burkus et al., 2003	Retrospective	rhBMP2 n=277	single-level	single-level primary	NR	rhBMP2 42±10	NR	rhBMP2 48.7	rhBMP2 175±36	Tobacco use rhBMP2	Other significant

Investigator (yr, country, ref #) Surgical Site	Study design	Comparison(s) No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD yrs (rng)	≥ 65 yrs (%)	Males (%)	Weight mean ± SD lbs (rng)	Comorbidities (%)	Comment
On-label use											
(Integrated analysis) USA **Lumbar spine** Note: may include pts in Burkus et al., 2003 (“Radio-graphic assessment …”) **KQ2, KQ3**	combined comparative analysis	(dose NR) ICBG n=402	lumbar DDD	anterior lumbar fusion with interbody fusion cages		 ICBG 41±10 p=0.007		 ICBG 52.2	 ICBG 179±38	31.4 Alcohol use rhBMP2 37.9 Tobacco use ICBG 32.8 Alcohol use ICBG 34.1	differences include previous back surgeries (lower in ICBG group), use of non-narcotic, weak narcotic, and muscle relaxant medications (all higher in rhBMP2 group)

Investigator (yr, country, ref #) Surgical Site	Study design	Comparison(s) No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD yrs (rng)	≥ 65 yrs (%)	Males (%)	Weight mean ± SD lbs (rng)	Comorbidities (%)	Comment
Off-label use											
Boden et al., (2002) USA **Lumbar Spine** **KQ2, KQ3**	Multi-center nonblinded RCT	rhBMP2/CRM plus Texas Scottish Rite Hospital (TSRH) Spinal System (TSRHSS) n=11	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 ICBG	grade I spondylo-listhesis	rhBMP2/CRM /TSRHSS 58±4	NR	rhBMP2/CRM /TSRHSS 27	NR	Tobacco use rhBMP2/CRM /TSRHSS 0 Alcohol use rhBMP2/CRM /TSRHSS 54 Diabetes rhBMP2/CRM /TSRHSS 0 Previous back surgery rhBMP2/TSRHSS 27%	Other than diabetes, no significant differences between groups
		(40 mg/pt) rhBMP2/CRM alone n=11				rhBMP2/CRM alone 52±6		rhBMP2/CRM alone 56		Tobacco use rhBMP2/CRM alone 12 Alcohol use rhBMP2/CRM alone 25 Diabetes rhBMP2/CRM alone 0 Previous back surgery rhBMP2 alone 12%	

		(40 mg/pt) ICBG plus TSRHSS n=5				ICBG/TSRHSS 53±10		ICBG/TSRHSS 40		Tobacco use ICBG/TSRHSS 20 Alcohol use ICBG/TSRHSS 40 Diabetes ICBG/TSRHSS 40 (p=0.036 for diabetes) Previous Surgery?	
Burkus et al., (2005) USA **Lumbar Spine** Note: includes all pts from Burkus et al., 2002, rec# 11510; same pts as Burkus et al., 2006, rec# 6640 **KQ2, KQ3**	Multicenter, nonblinded RCT	rhBMP2 n=79 (8-12 mg/pt) ICBG N=52	single-level lumbar DDD	primary single-level anterior lumbar fusion with a pair of threaded allograft cortical bone dowels (CBD) plus rhBMP2 or ICBG	grade I spondylo-listhesis	rhBMP2 40 ICBG 44	NR	rhBMP2 40 ICBG 36	rhBMP2 172 ICBG 173	Tobacco use rhBMP2 33 Previous back surgery rhBMP2 37 Tobacco use ICBG 33 Previous back surgery ICBG 33	No significant differences between groups

Dawson et al., 2009 USA **Lumbar spine** **KQ2, KQ3**	Multicenter nonblinded RCT	rhBMP2/CRM n=25 (12 mg/pt) ICBG n=21	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	grade I spondylolisthesis	rhBMP2/CRM 56 ICBG 57	NR	rhBMP2/CRM 40 ICBG 43	rhBMP2/CRM 176 ICBG 185	Tobacco use rhBMP2/CRM 24 ICBG 24 Previous back surgery rhBMP2/CRM 24 ICBG 29	Previous back surgery not at index level
Dimar et al., (2009) USA **Lumbar Spine** "Note [AHRQ]: contains pts in Glassman et al., 2007, rec# 4040; Dimar et al., 2006 rec# 5480; Glassman et al., 2005, rec# 8040" **KQ2, KQ3**	Multicenter nonblinded RCT	rhBMP2/CRM n=239 (40 mg/pt) ICBG n=224	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	grade I spondylo-listhesis	rhBMP2/CRM 53 (20-82) ICBG 52 (18-86)	NR	rhBMP2/CRM 45 ICBG 42	rhBMP2/CRM 187 (103-361) ICBG 189 (99-312)	Tobacco use rhBMP2/CRM 26 Alcohol use rhBMP2/CRM 38 Previous back surgery rhBMP2 30 Tobacco use ICBG 26 Alcohol use ICBG 35 Previous back surgery ICBG 28	No significant differences between groups
Glassman et al., (2008) USA	Multicenter nonblinde	rhBMP2 n=50 (dose not	single- or multi-level lumbar	single- or multi-level primary	Not reported	rhBMP2 69±6	NR all > 60	rhBMP2 30	NR BMI rhBMP2	Tobacco use rhBMP2 22	No significant differences

Lumbar Spine **KQ2, KQ3**	d RCT	reported) ICBG n=52	DDD	instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG		 ICBG 70±6		 ICBG 33	29±6 ICBG 28±6	 ICBG 17	between groups, including mean number of surgical levels (rhBMP2=1.96, ICBG=1.98)
Haid et al., (2004) USA **Lumbar Spine** **KQ2, KQ3**	Multicenter, nonblinded RCT	rhBMP2 n=34 (4.2-8.4 mg/pt) ICBG N=33	single-level lumbar DDD	single-level primary posterior lumbar interbody fusion (PLIF) with interbody fusion cages plus rhBMP2 or ICBG	grade I spondylo-listhesis	rhBMP2 46 (26-66) ICBG 46 (28-71)	NR	rhBMP2 50 ICBG 46	rhBMP2 180±38 ICBG 173±36	Tobacco use rhBMP2 53 Alcohol use rhBMP2 44 Previous back surgery rhBMP2 35 Tobacco use ICBG 46 Alcohol use ICBG 27 Previous back surgery ICBG 39	No significant differences between groups

Glassman et al., (2007) USA **Lumbar Spine** **KQ2, KQ3**	Retrospective with historical control group	rhBMP2 n=91 (12 mg/pt) ICBG n=35	single- and multi-level lumbar DDD, degenerative scoliosis, postdiscectomy instability, spinal stenosis, adjacent level degeneration	single- or multi-level primary or revision instrumented posterolateral lumbar fusion	Not reported	rhBMP2 60 (27-84) ICBG 53 (33-80)	NR	rhBMP2 40 ICBG 43	NR	Tobacco use rhBMP2 15 ICBG 23	No statistically significant differences between primary single-level pts in rhbMP2 or ICBG group
Mummaneni et al., 2004 USA **Lumbar Spine** **KQ2, KQ3**	Retrospective single-center cohort study	rhBMP2/AGB n=25 (8.4 mg/pt) ICBG N=19	single- or multi-level lumbar DDD	single- or multi-level primary transforaminal lumbar interbody fusion (TLIF) with interbody fusion cages with rhBMP2 plus AGB or ICBG alone	grade I spondylo-listhesis	rhBMP2/AGB 56±12 (33-76) ICBG 49±10 (33-64)	rhBMP2/AGB 24 ICBG 0 (p < 0.01)	rhBMP2/AGB 68 ICBG 47	NR	Tobacco use rhBMP2/AGB 12 Prior surgery rhBMP/AGB 40 Tobacco use ICBG 5 Prior surgery ICBG 67	More older pts and males in the rhBMP2/AGB group than ICBG group, but small numbers limit comparison
Pradhan et al., 2006 USA **Lumbar Spine**	Prospective consecutive patient single-	rhBMP2 n=9 (dose NR) ICBG	single-level lumbar DDD	single-level primary aAAnterior lumbar interbody	grade I spondylo-listhesis	rhBMP2 51 ICBG	3 (1 of 36)	rhBMP2 33 ICBG	NR	NR	Patient sample demographics not well described

KQ2, KQ3	center cohort study	n=27		fusion (ALIF) with femoral ring allograft (FRA) plus rhBMP2 or ICBG		53		18			
Singh et al., 2006 USA **Lumbar Spine** **KQ2, KQ3**	Prospective single-center case-matched cohort study	rhBMP2/ICBG n=39 (12-36 mg/pt) ICBG N=11	single- or multi-level lumbar DDD	single- or multi-level primary instrumented posterolateral lumbar fusion with rhBMP2 plus ICBG or ICBG alone	grade I-II spondylo-listhesis	rhBMP2/ICBG 65 ICBG 54	NR	rhBMP2/ICBG 44 ICBG 46	NR	NR	Patients in rhBMP2/ICBG group appear to be older, but no statistical analysis was done to confirm
Slosar et al., 2007 USA **Lumbar Spine** **KQ2, KQ3**	Prospective consecutive patient single-center cohort study	rhBMP2 n=45 (3-9 mg/pt) ALG N=30	single- or multi-level lumbar DDD	single- or multi-level primary instrumented anterior lumbar interbody fusion (ALIF) with femoral ring allograft (FRA) plus rhBMP2 or allograft bone chips (ALG)	grade I-II spondylo-listhesis	rhBMP2 45 ALG 44	NR	rhBMP2 60 ALG 51	NR	Tobacco use rhBMP2 18 Previous back surgery rhBMP2 46 Tobacco use ALG 8 Previous back surgery ALG 37	Both groups were similar in demographics and number of levels fused

Johnsson et al., 2002 Sweden **Lumbar Spine** **KQ2, KQ3**	Multicenter nonblinded RCT	rhBMP7 n=10 (7 mg/pt)	single-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	NR	rhBMP7 43±11	0	rhBMP7 30	NR	rhBMP7 40	Poorly described patients samples
		ICBG n=10				ICBG 40±10		ICBG 70		ICBG 30	
Kanayama et al., 2006 Japan, USA **Lumbar Spine** **KQ2, KQ3**	Multicenter nonblinded RCT	rhBMP7 n=9 (7 mg/pt)	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion with rhBMP7 or AGB/CRM	grade I spondylo-listhesis	rhBMP7 70±8	NR	rhBMP7 56	NR	NR	Poorly described patient samples, significantly older pts in rhBMP7 group
		AGB/CRM n=10				AGB/CRM 59±9 ($p < 0.05$)		AGB/CRM 60			
Vaccaro et al., 2008 USA **Lumbar Spine** **KQ2, KQ3**	Multicenter nonblinded RCT	rhBMP7 n=207 (7 mg/pt)	single-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	grade I-II spondylo-listhesis	rhBMP7 68±10	at least 50% in both groups rhBMP7 med=68	rhBMP7 34	NR NSD reported	NR	No significant differences between groups
		ICBG n=86				ICBG 69±8	ICBG med=71	ICBG 30			
Vaccaro et al., 2008 USA **Lumbar Spine** Note: Long-term F/U study that includes all pts	Multicenter, nonblinded RCT	rhBMP7 n=24 (7 mg/pt)	single-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	grade I-II spondylo-listhesis	rhBMP7 63 (43-80)	NR	rhBMP7 46	rhBMP7 198 (125-299)	NR	Patients in rhBMP7 group appear to be younger and heavier than in ICBG group, but no
		ICBG n=12				ICBG 67 (51-79)		ICBG 42	ICBG 176 (130-220)		

from Vaccaro et al., 2004, and Vaccaro et al., 2005 **KQ2, KQ3**											statistical analysis was done
Baskin et al., 2003 USA **Cervical Spine** **KQ2, KQ3**	Multicenter, nonblinded RCT	rhBMP2/ALG n=18 (0.6-1.2 mg/pt) ICBG/ALG n=15	single- or two-level cervical DDD	single- or two-level primary instrumented ACDF with rhBMP2/ALG or ICBG/ALG	NR	rhBMP2/ALG 51 ICBG/ALG 47	NR	rhBMP2/ALG 44 ICBG/ALG 47	rhBMP2/ALG 170 ICBG/ALG 174	Tobacco use rhBMP2/ALG 28 ICBG/ALG 47	No significant differences between groups
Buttermann et al., 2008 USA **Cervical Spine** **KQ2, KQ3**	Prospective nonrandomized cohorts of consecutive patients	rhBMP2/CRA n=30 (0.9-3.7 mg/pt) ICBG n=36	single- or multiple-level cervical DDD	single- or multi-level primary instrumented or uninstrumented ACDF with rhBMP2/CRA or ICBG	NR	rhBMP2/CRA 49±10 ICBG 48±9	NR	rhBMP2/CRA 50 ICBG 33	NR	Tobacco use rhBMP2/CRA 37 Adjacent level DDD rhBMP2 63 Tobacco use rhBMP2/CRA ICBG 53 Adjacent level DDD ICBG 64	No significant differences between pt groups except a greater number of levels were treated in the rhBMP2/CRA group compared to the ICBG group (mn 1.6 vs. 2.2, p=0.003)
Crawford et al., 2009 USA	Retrospective cohort of consecutiv	rhBMP2/BGE n=41 (4.2-12 mg/pt)	single- or multi-level posterior cervical	single- or multi-level instrumented posterior	NR	rhBMP2/BGE 56±11	NR	rhBMP2/BGE 32	NR	Tobacco use rhBMP2/BGE 24	No significant differences between

Cervical Spine **KQ2, KQ3**	e patients	ICBG n=36	stenosis, ACDF nonunion, or unstable spondylosis	cervical spinal fusion with rhBMP2/BGE or ICBG		ICBG 54±12		ICBG 42		ICBG 36	groups
Smucker et al., 2006 USA **Cervical Spine** **KQ2, KQ3**	Retrospective case-control	rhBMP2/CRA n=69 (dose NR)	NR	single- or multi-level instrumented ACDF with rhBMP2/CRA or CRA alone	NR	rhBMP2/CRA 52	NR	rhBMP2/CRA 49	NR	Tobacco use rhBMP2/CRA 29 Prior ACDF rhBMP2/CRA 28 ≥ 3 levels fused rhBMP2/CRA 13	Patients in rhBMP2/CRA (cortical ring allograft) group had significantly higher rates of comorbidities that can adversely affect fusion
		CRA n=165				CRA 50		CRA 49		Tobacco use CRA 14 (p=0.02) Prior ACDF CRA 10 (p=0.001) ≥ 3 levels fused CRA 2 (p=0.003)	
Vaidya et al., 2007 USA **Cervical Spine** **KQ2, KQ3**	Retrospective cohort of consecutive patients	rhBMP2 n=22 (1-3 mg/pt)	single- or multiple-level cervical DDD	single- or multi-level primary instrumented ACDF with interbody	NR	rhBMP2 50 (29-70)	NR	rhBMP2 32	NR	NR	No significant differences between groups
		ALG/DBM n=24				ALG/DBM 48 (30-69)		ALG/DBM 45			

				fusion cages rhBMP2 on ACS or ALG/DBM							

Appendix Table 4. Comparative studies reported after the AHRQ HTA search period evaluating BMPs in spinal fusion: patient demographics.

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
On-label use										
FDA SSED: InFUSE (P000058) **KQ3** (overlaps with Boden 2000, Burkus 2002, Burkus 2003)	Integrated analysis (of pilot (Boden 2000[13]) and pivotal (Burkus 2002[14] + subset of Burkus 2003[15]))	rhBMP-2: n = 288	Single-level DDD	single-level primary anterior open or laproscopic (n = 134 BMP pts only) lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	NR	42.86 (19.2–78.4)*	48.7% (135/277)*	174.6 lbs	Tobacco use: 31.4% (87/277)* Previous back surgery: 31.4% (87/277)*	**Randomization:** patients who underwent fusion via the open surgical approach were randomized (and reported in Burkus 2002[14]); those who underwent fusion via the laproscopic approach were not randomized (non-randomized investigational arm).
		ICBG: n = 139				42.3 (19.0–70.6)*	50% (68/136)*	181.1 lbs	Tobacco use: 36.0% (49/136)* Previous back surgery: 40.4% (55/136)*	

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
Burkus et al. (2011) USA **Lumbar spine** (patients from FDA	Cohort study: integrated analysis of 3 studies	BMP-2: n = 1093 (varying surgical interventions) (dose NR)	Single-level symptomatic DDD, grade I spondylolisthesis or lower, or disabling back and/or leg pain	Study #1 (on-label use) (patients from FDA SSED Pivotal Study; also reported in Burkus 2002 and subset of Burkus 2003 integrated	NR	NR	NR	NR	NR	No demographic details were reported

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
SSED Pivotal Study; also reported in Burkus 2002 and subset of Burkus 2003 integrated analysis, Dimar 2009 RCT, as well as from Gornet 2007 RCT (abstract only))		Autograft (ICBG): n = 360	(further details NR)	analysis): ALIF with LT-CAGE done laproscopically (n = 134, BMP only, nonrandomized arm) or with open surgery (BMP2, n = 143; ICBG, n = 136, randomized arm) Study #2 (Gornet 2007 RCT): open ALIF with BMP (all pts) using lumbar tapered fusion device (n = 172) (on-label use) or metal-on-metal lumbar disc arthroplaty device (n = 405) (off-label use). Study #3 (off-label use) (Dimar 2009): single-level instrumented posterolateral lumbar arthrodesis through open						

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
				approach with BMP-2-matrix (n = 239) or ICBG (n = 224).						
Carragee et al. (2011) USA **Lumbar spine** **KQ3**	Retrospective cohort	rhBMP2: n = 69 (4.2 mg/pt)	rhBMP2: Deg. Spondylo: n = 33 (48%) Low-grade isthmic spondylo: n = 23 (33%)	1- or 2-level ALIF including L5/S1 via an open retroperitoneal approach with a femoral ring allograft or titanium mesh	rhBMP2 1-level: n = 45 (65%) 2-level: n = 24 (35%)	rhBMP2: 42.4 ± 10.3 (range, 22–65)	100%	rhBMP2: 81 ± 12.1 kg	NR	Groups well matched for age, diagnosis, number of levels fused, weight. Data from a prospective database of

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
			Recurrent herniation/ DDD: n = 13 (19%)	cage filled with ICBG or rhBMP-2/ACS						consecutive patients, retrospectively analyzed
		Osteophytes or ICBG: n = 174	ICBG: Deg. Spondylo: n = 80 (46%) Low-grade isthmic spondylo: n = 54 (31%) Recurrent herniation/ DDD: n = 40 (23%)		ICBG: 1-level: n = 110 (59%) 2-level: n = 64 (41%)	ICBG: 40.9 ± 9.9 (range, 25–65)		ICBG: 79 ± 13.4 kg		
Crawford et al. (2010) USA **Sacrum**	Retrospective cohort	rhBMP2: n = 39 (dose NR)	Idiopathic scoliosis	Posterior extension of an existing fusion to the sacrum with segmental pedicle screw	rhBMP2: previous levels fused: 9.9 ± 2.7; new levels fused: 2.6 ± 1.7	rhBMP2: 49.8 ± 10.5 years	rhBMP2:8.3	NR	NR	Data from a single institution, prospective database, retrospectively analyzed;

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
KQ2, KQ3 (appears to contain the same patients reported in Maeda (2009))		Autogenous graft (historical controls): N = 25		instrumentation, including S1 pedicle screw fixation and iliac screw fixation; all but five patients (study group) had anterior interbody device support at the lowest level via an anterior or transforaminal approach.	Autogenous graft: previous levels fused: 10.2 ± 2.2; new levels fused: 2.6 ± 1.8	Autogenous graft: 43.5 ± 10.2	Autogenous graft: 4.2			radiographs were analyzed retrospectively Use of historical controls Groups were well matched with respect to demographic, radiographic, and surgical data with the following exceptions: the control group was younger (43.5 *vs.* 49.8 years; *P* = .04), had more anterior levels fused (3.3 *vs.* 1.7; *P* = .01), and more thoraco-abdominal approaches (25% *vs.* 2.7%; *P* = .01)
Howard et al. (2011) USA **Lumbar spine** **KQ3**	Cross-sectional	rhBMP2: n = 59 (dose NR) ICBG: n = 53	NR	1- to 2- level instrumented posterolateral fusion from L1 to S1	NR	Overall: 56.6 years (range, 16–84) NR by treatment group	Overall: 35.7% (40/112) NR by treatment group	NR	NR	
Joseph et al.	Prospective	rhBMP2:	Spondylo-	Minimal access	1-level fusion:	49.7 (22-	61%	NR	NR	Demographics NR

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
(2007) [32 32 31 30] USA **Lumbar spine** **KQ3**	cohort	n = 23 (24 levels) (4.2 mg/level) Local autograft: n = 10 (12 levels)	listhesis: 85% (28/33) DDD: 15% (5/33)	PLIF or TLIF with interbody cages and percutaneous pedicle screw fixation	91% (30/33) 2-level fusion: 9% (3/33)	69) years	(20/33)			separately for BMP-2 vs. autograft groups.
Latzman et al. (2010) USA **Lumbar spine** **KQ3**	Retrospective cohort	rhBMP2: n = 24† (12 mg/8 cc; 24 mg/16 cc) Auto- or allograft only: N = 105†	NR	Lumbar and lumbosacral spinal fusion with and without interbody cage placement	NR	rhBMP2: 50.1 ± 12.7 years‡ Auto- or allograft only: 55.8 ± 11.5 years‡	rhBMP2: 77.7‡ Auto- or allograft only: 89.8‡	NR	rhBMP2: diabetes, 7.4 current smoker, 44.4 Auto- or allograft only: diabetes, 18.5 current smoker, 39.8	Retrospective chart review Significantly more patients in the BMP group had interbody cage placement: 70% vs. 30%, *P* = .001; and received allograft without autograft: 52% vs. 22%, *P* = .002
Lee et al. (2010) USA **Lumbar spine** **KQ2, KQ3**	Retrospective cohort	rhBMP2 (with allograft): n = 86 (4.2 mg/2/8 ml for 1-level; 8.4 mg/5.6 ml for 2-level; 12 mg/8.0 ml for 3+ levels) age ≥ 65 years: n = 34 age < 65 years: n = 52	DDD	Instrumented posterior lumbar fusion	rhBMP2 age ≥ 65 years: 1-level fusion: 50.0 2-level fusion: 50.0 Revision: 35.3 rhBMP2 age < 65 years: 1-level fusion: 75.0 2-level fusion: 25.0	rhBMP2 age ≥ 65 years: 74.1 ± 5.8 years (65–91) rhBMP2 age < 65 years: 49.9 ± 11.2 (17–64)	rhBMP2 age ≥ 65 years: 52.9% rhBMP2 age < 65 years: 38.5%	NR	rhBMP2 age ≥ 65 years: medical comorbidity: 52.9 osteoporosis: 41.2 smoking: 14.7 rhBMP2 age < 65 years: medical comorbidity: 17.3	All the patients with osteoporosis were taking bisphosphonates during the postop follow-up period; Smokers = smoked continuously for at ≥ 1 year prior to surgery, as well as postoperatively; Medical comorbidities =

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
					Revision: 50.0				osteoporosis: 11.5 smoking: 26.9	patients receiving treatment for two or more concurrent medical diseases of Li et al's comorbidity definition, such as diabetes mellitus, hypertension and thyroid disease, etc.; Revision surgeries were restricted to cases in which surgery was performed for pseudoarthrosis.
		ICBG: age ≥ 65 years: n = 41			ICBG age ≥ 65 years: 1-level fusion: 31.7 2-level fusion: 68.3 Revision: 19.5	ICBG age ≥ 65 years: 72.4 ± 5.1 (65–83)	ICBG age ≥ 65 years: 41.5%		ICBG age ≥ 65 years: Medical comorbidity: 58.5 osteoporosis: 43.9 smoking: 17.1	
Rihn et al. (2009) [35 35 34 33] USA **Lumbar spine** **KQ3**	Retrospective cohort	rhBMP2 n = 86 ICBG n = 33	DDD: 10.9% DDD/HNP: 12.6% Recurrent HNP: 27.7% Isthmic spondylo-listhesis: 32.8% Degener-ative spon-dylolisthesis: 15.1% Failed laminectomy and fusion: 0.8%	TLIF 1-level	Levels per patient: 1	47.4 years	52.9%	NR	Previous lumbar surgery: 37.0%	Demographic data not reported separately for each treatment group.
Taghavi et al. (2010)	Retrospective cohort	rhBMP2 n = 24 (1.5 mg/mL	symptomatic pseudarthro sis (pain	Transpedicular, instrumented revision	rhBMP2 Levels per patients: 2.0	rhBMP2 57.3 ± 11.6	rhBMP2: 45.8	NR	rhBMP2 smokers: 8.3 diabetes: 8.3	

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
USA **Lumbar spine** **KQ2, KQ3**		concentration; 12 mg)	and/or instability) following a previous PLF for degenerative conditions of the lumbar spine, such as degenerative disc disease, stenosis, or spondylolisthesis.	posterolateral fusion; rhBMP2 (INFUSE kit, 12 mg, 1.5 mg/mL concentration, ACS); BMMA from a single iliac crest; autograft	1-level: 54.2% 2-level: 16.7% 3-level: 16.7% 4-7 levels: 12.5%	years (31–75)			osteoporosis: 12.5	
		BMAA n = 18			BMAA Levels per patient: 2.2 1-level: 38.9% 2-level: 33.3% 3-level: 16.7% 4-7 levels: 11.0%	BMAA 59.7 ± 11.6 years (40–77)	BMAA 55.6		BMAA smokers: 11.1 diabetes: 5.5 osteoporosis: 11.1	
		Autograft n = 20			Autograft Levels per patient: 1.9 1-level: 50.0% 2-level: 25.0% 3-level: 15.0% 4-7 levels: 10.0%	Autograft 55.8 ± 13.2 years (21–73)	Autograft 55.0		Autograft smokers: 15.0 diabetes: 10.0 osteoporosis: 10.0	
Vaidya, Weir et al. (2007) [37 37 36 35] USA **Lumbar and cervical spine** **KQ3**	Prospective cohort	rhBMP2 + allograft: n = 36 (55 levels) (2 mg/level for lumbar; 1 mg/level for cervical))	rhBMP2 + allograft: Adult scoliosis: 19% (7/36) Revision lumbar surgery: 25% (9/36) Spondylo-listhesis: 11% (4/36)	rhBMP2 + allograft: ALIF: n = 13 (20 levels) TLIF: n = 12 (17 levels) Anterior cervical decompression/ fusion: n = 11 (18 levels)	NR	rhBMP2 + allograft: 47.9 (18-71) years	rhBMP2 + allograft: 56% (20/36)	NR	NR	Demographic NR separately for lumbar vs cervical AHRQ excluded this study (as “not relevant design”), so will be evaluated for KQ3 only.

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
			Discogenic pain: 14% (5/36) Cervical disc herniation: 25% (9/36) Cervical myelopathy: (2/36)							
		DBM + allograft: n = 41 (63 levels)	DBM + allograft: Adult scoliosis: 12% (5/41) Revision lumbar surgery: 32% (13/41) Spondylo-listhesis: 10% (4/41) Discogenic pain: 17% (7/41) Cervical disc herniation: 22% (9/41)	DBM + allograft: ALIF: n = 11 (16 levels) TLIF: n = 18 (25 levels) Anterior cervical decompression/ fusion: n = 12 (22 levels)		DBM + allograft: 45 (16-77) years	DBM + allograft: 44% (18/41)			

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
			Cervical myelopathy: 7% (3/41)							
Delawi et al. (2010) Europe (Netherlands, France, Italy, Spain) **Lumbar spine** **KQ2, KQ3**	RCT Multicenter (5)	OP-1 (rhBMP-7): n = 18 (3.5 mg per side of spine)	rhOP-1: Deg. Spondylo: n = 10 (56%) Isthmic spondylo: n = 8 (44%)	1-level, posterolateral lumbar fusion using pedicle screw instrumentation; decompression via bilateral laminectomy or partial laminectomy and medial facetectomy; under general anesthesia; prophylactic cephalosporin given for 24 hours starting 15 mins. before incision	rhOP-1 L3-L4: 22% L4-L5: 28% L5-L6: 0% L5-S1: 50%	rhOP-1 53 ± 18 years	rhOP-1 55.6%	rhOP-1 26 ± 4 kg/m^2 Autograft: 27 ± 3 kg/m^2	rhOP-1 Smoker: 44.4%	Surgical techniques strictly standardized and identical for the 2 groups with the exception of the bone grafting technique
		Autograft (ICBG): n = 16	Autograft: Deg. Spondylo: n = 11 (69%) Isthmic spondylo: n = 5 (31%)		Autograft L3-L4: 13% L4-L5: 75% L5-L6: 6% L5-S1: 6%	Autograft 55 ± 13 years	Autograft: 37.5%		Autograft Smoker: 25.0%	
Hwang et al. (2010) USA; Canada **Lumbar spine** **KQ3** (same	RCT Multicenter (24)	rhOP-1: Initial phase: n = 208 Extension phase: n = 144*	rhOP-1: Deg. Spondylo grade I: n = 135 (93.8%)*; Deg. Spondylo grade II: n = 5 (3.5%)*; Unable to distinguish	Single-level decompression and uninstrumented posterolateral fusion of the listhetic segment	rhOP-1 angular motion (n = 138)*: mean 4.1°; translational movement (n = 136): mean 1.8 mm levels fused*: L3-4 11.8%; L4-5 86.1%; L5-S1 2.1%	Overall: 68 years (36–84)*	rhOP-1: 34.7%*	NR	NR	No statistically significant differences between treatment groups were noted Majority of patients in each group were older than 65 years of age

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
patients as Vaccaro, Lawrence (2008)[44]; additional safety data reported)			b/w grade I and II: n = 4 (2.8%)*							
		Autograft: Initial phase: n = 87 Extension phase: N = 58*	Autograft: Deg. Spondylo grade I: n = 54 (93.1%)*; Deg. Spondylo grade II: n = 2 (3.4%)*; Unable to distinguish b/w grade I and II: n = 2 (3.4%)*		Autograft angular motion (n = 51)*: mean 4.3°; translational movement (n = 51)*: mean 1.5 mm; levels fused*: L3-4 15.5%; L4-5 82.8%; L5-S1 1.7%		Autograft: 27.6%*			
Xu et al. (2011) USA **Cervical spine** **KQ2, KQ3**	Retrospective cohort study	rhBMP-2 + some/all of the following (DBM (31%), local autograft (77%), allograft (21%), hydroxyapatite crystals (61%): n = 48 (dose NR)	Symptomatic primary degenerative cervical pathologies (no further diagnoses were reported) rhBMP2: Back pain: 73% (35/48) Radiculop.: 39% (18/48) Motor weakness: 72% (33/48)	Primary posterior cervical arthrodesis	rhBMP2: levels fused: 6.3 ± 2.2	rhBMP2: 60.3 ± 15.0 years	rhBMP2: 48% (23/48)	NR	rhBMP2: diabetes: 15% (8/48) CAD: 11% (6/48) Osteoporosis: 9% (4/48) Obesity: 9% (5/48) Smoking history: 30% (16/48) Hypertension: 47% (25/48) Previous surgery: 30% (16/48)	No statistically significant differences between treatment groups were noted

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
		Non-BMP: some/all of the following (DBM (86%), local autograft (88%), allograft (72%), hydroxyapatite crystals (0%): n = 156	Sensory deficits: 54% (25/48) Bowel/bladder dysfuction: 18% (8/48) Non-BMP: Back pain: 64% (99/156) Radiculop.: 41% (64/156) Motor weakness: 80% (124/156) Sensory deficits: 55% (85/156) Bowel/bladder dysfunction: 23% (36/156)		Non-BMP: levels fused: 5.8 ± 1.8	Non-BMP: 60.8 ± 12.7 years	Non-BMP: 64% (100/156) (*P* = .05)		Non-BMP: diabetes: 25% (39/156) CAD: 15% (23/156) Osteoporosis: 3% (5/156) Obesity: 12% (18/156) Smoking history: 22% (35/156) Hypertension: 54% (85/156) Previous surgery: 27% (42/156)	
Yaremchuk (2010) USA **Cervical spine** **KQ2, KQ3**	Retrospective cohort study	BMP (n = 260) -dosages NR Non-BMP (n = 515)	NR	Cervical spinal fusion with or without BMP	NR	NR	NR	NR	NR	Demographic data NR

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
Cahill (2011) USA **Lumbar spine** **KQ2, KQ3**	Retrospective case control (database) study	BMP (rhBMP-2 OR rhBMP-7) ± autograft ± allograft n = 2,372 (6% received autograft harvested from a different incision) (28% received allograft)	Lumbar disc herniation: 47% (1104/2372) DDD: 64% (1507/2372) Spondylosis: 22% (528/2372) Spinal stenosis: 40% (946/2372) Spondylolisthesis: 34% (814/2372) Back pain: 28% (670/2372)	Single-level lumbar fusion (any approach) with or without BMP (rhBMP-2 or rhBMP-7). Fusion type: Interbody: 35% Posterolateral: 18% Circumeferential: 48% Instrumented fusion: 87% (%s similar in both groups)	NR (All patients received single-level fusion)	48 years	51%	NR	Charlson comborbidity score (mean, median): 0.3 (0) Osteoporosis: 1% (19/2372) Tobacco use: 27% (633/2372) Obesity: 14% (326/2732) Diabetes: 11% (268/2372)	No significant differences in any baseline characteristics between groups.
		No BMP ± autograft ± allograft (matched controls): n = 2,372 (19% received autograft	Lumbar disc herniation: 44% (1055/2372) DDD: 63% (1501/2372) Spondylosis: 23% (544/2372)			48 years	49%		Charlson comborbidity score (mean, median): 0.3 (0) Osteoporosis: 1% (17/2372)	

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
		harvested from a different incision) (26% received alloigraft) Propensity scores used to match patients that underwent fusion with BMP to controls with a similar probability of undergoing a fusion with BMP.	Spinal stenosis: 40% (960/2372) Spondylolisthesis: 36% (844/2372) Back pain: 29% (686/2372)						Tobacco use: 26% (613/2372) Obesity: 14% (329/2732) Diabetes: 10% (248/2372)	
Deyo et al. (2011) USA **Lumbar spine** **KQ3**	**Retrospective cohort (database) study**	BMP: n = 1703	Spinal stenosis: 100% Spondylolisthesis: 42.3% (721/1703) Scoliosis: 12.1% (206/1703)	<u>Fusion type:</u> Simple fusion: 59.5% (1014/1703) Complex fusion: 40.5% (689/1703) <u>Levels fused:</u> 1-2: 61.7% (1050/1703) 3+: 17.4% (296/1703) Unknown: 21.0%	See diagnosis	74.9 ± 4.7 years	34.6% (589/1703)	NR	Previous spine surgery: 21.5% (366/1703) Quan comorbidity score: 0: 49.4% (841/1703) 1+: 50.6% (862/1703)	Significant differences in age, number of levels fused, fusion type, and history of spinal surgery. Regression analysis performed to adjust for these variables and did not affect the results.

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
									Diabetes: 21.5% (366/1703) Smoker: 5.1% (87/1703)	
		No BMP: n = 15,119	Spinal stenosis: 100% Spondylo-listhesis: 45.1% (6814/15,119) (*P* = .03) Scoliosis: 11.0% (1657/15,119) (*P* = .16)	<u>Fusion type</u>: (*P* < .001) Simple fusion: 71.4% (10,792/15,119) Complex fusion: 28.6% (4327/15,119) <u>Levels fused:</u> (*P* < .001) 1-2: 41.9% (6330/15,119) 3+: 10.8% (1635/15,119) Unknown: 47.3%		75.3 ± 4.9 years *P* = .001	34.6% (5233/15,119) *P* = .98		Previous spine surgery: 14.4% (2181/15,119) (*P* < .001) Quan comorbidity score: (*P* = .27) 0: 50.8% (7681/15,119) 1+: 49.2% (7438/15,119) Diabetes: 20.2% (3054/15,119) (*P* = .21) Smoker: 4.9% (744/15,119) (*P* = .74)	
Mines et al. (2011) USA **Lumbar spine**	Retrospective cohort (database) study	BMP: n = 15,640	NR	Lumbar fusion surgery with or without BMP (BMP2 or BMP7)	NR	74.2 ± 5.1 years	33.0% (5102/15,460)	NR	Diabetes: 36.4% (5625/15,460) Chronic pancreatitis: 0.9%	Significant differences in age, age group (NR here), gender, race (NR here), diabetes, and previous

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
KQ3									(140/15,460)	cholecystectomy. Regression analysis performed.
									Alcoholism: 1.5% (225/15,460)	
									Cholecyst-ectomy: 3.5% (539/15,460)	
		No BMP: n = 78,194				74.6 ± 5.2 years (*P* < .001)	34.6% (27,071/ 78,194) (*P* < .001)		Diabetes: 35.5% (27,777/78,194) (*P* = .041)	
									Chronic pancreatitis: 1.0% (744/78,194) (*P* = .590)	
									Alcoholism: 1.5% (225/78,194) (*P* = .383)	
									Cholecyst-ectomy: 3.0% (2321/78,194) (*P* < .001)	
Cahill et al. (2009) USA **Lumbar,**	Retrospective cohort (database) study	BMP: n = 17,623	<u>BMP:</u> DDD or disc herniation: 70.72% (12/463/17,623)	Fusion (any) <u>BMP:</u> Revision fusion: 8.52% (1502/17,623)	NR	53.79 ± 14.07	43.74% (7708/17,623)	NR	<u>BMP:</u> Charlson Comorbidity Score: 0.48 ± 0.89	Disparities between groups: Women more likely to receive BMP than men (OR: 1.12 (95%

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
cervical, or thoracic spine **KQ3**			Other: 29.28% (5160/17,623)	Cervical: 16.38% (2886/17,623) Lumbosacral: 79.28% (13,972/17,623) Thoracolumbar: 4.23% (746/17,623) Unknown: 0.11% (19/17,623) Vertebral levels, 2-3: 83.03% (14,633/17,623) Vertebral levels, ≥4: 16.97% (2990/17,623)					CI, 1.09,1.16); nonwhite patients less likely to receive BMP than white patients (OR: 0.80 (95% CI, 0.75, 0.85); use of BMP decreased with increasing medical comorbidities according to Charlson socre (incremental OR: 0.95 (95% CI, 0.93, 0.96) per unit increase); conditions other than DDD or disc herniation associated with increased BMP use (OR: 1.28 (95% CI, 1.23, 1.33); BMP more commonly used in revision procedures (OR: 1.81 (95% CI, 1.69, 1.93); between percentages of patients in BMP vs no BMP group who underwent cervical fusion (16.38% vs.
		No BMP: n = 53,026	<u>No BMP:</u> DDD or disc herniation: 75.65% (40,116/53,026) Other: 24.35% (12,910/53,026)	<u>No BMP:</u> Revision fusion: 4.89% (2595/53,026) Cervical: 52.03% (27,589/53,026) Lumbosacral: 43.06% (22,8354.74%		53.26 ± 13.91	46.65% (24,738/53,026)		<u>No BMP:</u> Charlson Comorbidity Score: 0.53 ± 1.10	

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
				(2511/53,026) Thoracolumbar4 .74% (2511/53,026) Unknown: 0.17% (91/53,026) Vertebral levels, 2-3: 84.57% (44,846/53,026) Vertebral levels, ≥4: 15.43% (8180/53,026)						52.03%) and lumbosacral fusion (79.28% vs. 43.06%).

Investigator (yr, country) Surgical site	Study design	Comparison(s) # patients (n) (BMP dose)	Patient diagnosis	Surgical intervention	Defect severity and characteristics (%)	Age mean ± SD (range)	% males	Weight mean ± SD (range)	Comorbidities (%)	Comment
Off-label use										
Yaremchuk (2010) **USA** Retrospective cohort study **Cervical spine**	Retrospective cohort study	BMP (n = 260) non-BMP (n = 515)	NR	Spinal fusion with or without BMP	NR	NR	NR	NR	**NR**	Type of BMP used not specified

ACS: absorbable collagen sponge; ALIF: anterior lumbar interbody fusion; BMAA: bone marrow aspirate with allograft; CAD: coronary artery disease; DDD: degenerative disc disease; FDA: Food and Drug Adminstration; f/u: follow-up; HNP: herniated nucleus pulposus; ICBG: iliac crest bone graft; IDE: investigational device exemption; IQR: interquartile range; ODI: Oswestry Low Back Pain Disability Questionnaire; PLF: posterolateral lumbar fusion; PLIF: posterior interbody fusion; RCT: randomized controlled trial; rhBMP2: recombinant human bone morphogenetic protein 2; SF-36: Short-Form 36; Spondylo: spondylolisthesis.

*FDA SSED for InFUSE: demographics reported for 277/288 investigational and 136/139 control patients; ie., demographics reported for patients in the pivotal but not the pilot portion of the population.
† Lumbar and lumbosacral fusion was performed on 125 patients; 101 patients underwent 104 operations without rhBMP2 and 20 underwent 23 operations with rhBMP2. Four patients had 1 operation with rhBMP2 and 1 without rhBMP2, for a total of 8 operations. There were 135 total operations.
‡ Age, percent male and comorbidities based on the number of operations: rhBMP2, n = 27; auto/allograft, n = 108.

Appendix Table 5. Comparative studies reported in the AHRQ HTA evaluating BMPs in spinal fusion: perioperative outcomes

Note. Abstraction tables copied directly from the AHRQ HTA report except that the references were changed to correspond to the current report. In addition, adverse events and complications were omitted as they were reported elsewhere.

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Mean OR time (hr)	Mean estimated blood loss (mL)	Mean hospital LOS (days)	Comment
On-label use								
Boden et al., 2000 USA **Lumbar spine**	Multicenter, nonblinded RCT	rhBMP2 (4.2-8.4 mg/pt) n=11	single-level DDD	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	rhBMP2 1.9±0.2 (2.3-4.2)	rhBMP2 95±31 (25-400)	rhBMP2 2.0±0.6 (0-6)	Besides OR time, no other significant differences reported
		ICBG n=3			ICBG 3.3±0.6 (1.0-3.2) p=0.006	ICBG 167±117 (50-400)	ICBG 3.3±1.4 (1-6)	
Burkus et al., 2002 USA **Lumbar spine**	Multicenter, nonblinded RCT	rhBMP2 (4.2-8.4 mg/pt) n=143	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	rhBMP2 1.6	rhBMP2 110	rhBMP2 3.1	No significant differences reported
		ICBG n=136			ICBG 2.0	ICBG 153	ICBG 3.3	
Burkus et al., 2003 (Integrated analysis) **Lumbar spine** Note: may include pts	Retrospective combined comparative analysis	rhBMP2 n=277 (dose NR)	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages	rhBMP2 1.8±0.8	rhBMP2 127±295	rhBMP2 2.2±1.7	Significantly more reoperations were reported in ICBG group than rhBMP2 group (p=0.0036)
		ICBG n=402			ICBG 2.7±1.3 p< 0.001	ICBG 193±414 p=0.024	ICBG 3.1±3.2 p < 0.001	

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Mean OR time (hr)	Mean estimated blood loss (mL)	Mean hospital LOS (days)	Comment
in Burkus et al., 2003 ("Radio-graphic assessment …")								
Off-label use								
Boden et al., (2002) USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP2/CRM plus Texas Scottish Rite Hospital (TSRH) Spinal System (TSRHSS) n=11	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 ICBG	rhBMP2/CRM /TSRHSS 3.7±0.3	rhBMP2/CRM /TSRHSS 577±113	rhBMP2/CRM /TSRHSS 3.3±0.1	No significant intergroup differences other than mean OR time
		(40 mg/pt) rhBMP2/CRM alone n=11			rhBMP2/CRM alone 2.0±0.2	rhBMP2/CRM alone 333±121	rhBMP2/CRM alone 4.0±0.9	
		(40 mg/pt) ICBG plus TSRHSS n=5			ICBG/TSRHSS 3.1±0.4 (p=0.002 rhBMP2/CRM alone vs other 2 groups)	ICBG/TSRHSS 430±82	ICBG/TSRHSS 4.4±0.5	
Burkus et al., (2005) USA	Multicenter, nonblinded RCT	rhBMP2 n=79 (8-12 mg/pt)	single-level lumbar lumbar DDD	primary single-level anterior	rhBMP2 1.4	rhBMP2 87	rhBMP2 2.9	Perioperative outcomes were

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Mean OR time (hr)	Mean estimated blood loss (mL)	Mean hospital LOS (days)	Comment
Lumbar Spine Note: includes all pts from Burkus et al., 2002, rec# 11510; same pts as Burkus et al., 2006, rec# 6640		ICBG N=52		lumbar fusion with a pair of threaded allograft cortical bone dowels (CBD) plus rhBMP2 or ICBG	ICBG 1.9 (p < 0.001)	ICBG 185 (p < 0.001)	ICBG 3.3 (p=0.20)	significantly better in the rhBMP2 group than the ICBG group
Dawson et al., 2009 USA **Lumbar spine**	Multicenter nonblinded RCT	rhBMP2/CRM n=25 (12 mg/pt)	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	rhBMP2/CRM 2.4±0.7 (95% CI, 2.1, 2.7)	rhBMP2/CRM 329±212 (95% CI, 241, 417)	rhBMP2/CRM 4.0±1.4 (95% CI, 3.4, 4.6)	No significant differences reported between groups
		ICBG n=21			ICBG 2.8±0.8 (95% CI, 2.2, 3.0)	ICBG 452±210 (95% CI, 357, 548)	ICBG 4.1±1.1 (95% CI, 3.6, 4.6)	
Dimar et al., (2009) USA **Lumbar Spine** Note: contains pts in Glassman et al., 2007,	Multicenter nonblinded RCT	rhBMP2/CRM n=239 (40 mg/pt)	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	rhBMP2/CRM 2.5±0.09	rhBMP2/CRM 343±265	rhBMP2/CRM 4.1±2.3	No surgical reintervention was related to recurrent stenosis or inadequate decompression
		ICBG n=224			ICBG 2.9±1.0 (p < 0.001)	ICBG 449±302 (p < 0.001)	ICBG 4.0±1.9	

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Mean OR time (hr)	Mean estimated blood loss (mL)	Mean hospital LOS (days)	Comment
rec# 4040; Dimar et al., 2006 rec# 5480; Glassman et al., 2005, rec# 8040								
Glassman et al., (2008) USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP2/ACS n=50 (dose not reported) ICBG n=52	single- or multi-level lumbar DDD	single- or multi-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	rhBMP2 4.1±0.6 ICBG 4.5±1.0 (p=0.024)	rhBMP2 670±487 ICBG 675±456	NR	Bone graft filler/extender used in 100% rhBMP2 and 67% ICBG cases, available local bone used in all cases
Haid et al., (2004) USA **Lumbar Spine**	Multicenter, nonblinded RCT	rhBMP2 n=34 (4.2-8.4) ICBG N=33	single-level lumbar DDD	single-level primary posterior lumbar interbody fusion (PLIF) interbody fusion cages plus rhBMP2 or ICBG	rhBMP2 2.6 ICBG 3.0	rhBMP2 323 ICBG 373	rhBMP2 3.4 ICBG 5.2 (p=0.065)	No significant differences between pt groups
Glassman et al., (2007) USA **Lumbar Spine**	Retrospective with historical control group	rhBMP2 n=91 (12 mg/pt) ICBG N=35	single- and multi-level lumbar DDD, degenerative scoliosis, postdiscecto	single- or multi-level primary or revision instrumented posterolateral	3.2 (1.5-6)	542 (100-3,600)	NR	No significant differences noted **Outcomes corrected by**

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Mean OR time (hr)	Mean estimated blood loss (mL)	Mean hospital LOS (days)	Comment
			my instability, spinal stenosis, adjacent level degeneration	lumbar fusion				**Spectrum.**
Mumma-neni et al., 2004 USA **Lumbar Spine**	Retrospective single-center cohort study	rhBMP2/AGB n=25 (8.4 mg/pt) ICBG N=19	single- or multi-level lumbar DDD	single- or multi-level primary transforaminal lumbar interbody fusion (TLIF) with interbody fusion cages with rhBMP2 plus AGB or ICBG alone	NR	NR	NR	
Pradhan et al., 2006 USA **Lumbar Spine**	Prospective consecutive patient single-center cohort study	rhBMP2 n=9 (dose NR) ICBG n=27	single-level lumbar DDD	single-level primary anterior lumbar interbody fusion (ALIF) with femoral ring allograft (FRA) plus rhBMP2 or ICBG	NR	NR	NR	Salvage posterior fusions performed secondary to subsequent pseudarthrosis and intractable symptoms
Singh et al., 2006 USA	Prospective single-center case-matched	rhBMP2/ICBG n=39 (12-36 mg/pt)	single- or multi-level lumbar DDD	single- or multi-level primary instrumented	NR	NR	NR	

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Mean OR time (hr)	Mean estimated blood loss (mL)	Mean hospital LOS (days)	Comment
Lumbar Spine	cohort study	ICBG N=11		posterolateral lumbar fusion with rhBMP2 plus ICBG or ICBG alone				
Slosar et al., 2007 USA **Lumbar Spine**	Prospective consecutive patient single-center cohort study	rhBMP2 n=45 (3-9 mg/pt) ALG N=30	single- or multi-level lumbar DDD	single- or multi-level primary instrumented anterior lumbar interbody fusion (ALIF) with femoral ring allograft (FRA) plus rhBMP2 or allograft bone chips (ALG)	NR	NR	NR	Salvage posterior fusions performed secondary to subsequent pseudarthrosis
Johnsson et al., 2002 Sweden **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP7 n=10 (7 mg/pt) ICBG n=10	single-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	NR	NR	NR	No perioperative results reported
Kanayama et al., 2006 Japan, USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP7 n=9 (7 mg/pt) AGB/CRM n=10	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion with rhBMP7	NR	NR	NR	No perioperative results reported

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Mean OR time (hr)	Mean estimated blood loss (mL)	Mean hospital LOS (days)	Comment
				or AGB/CRM				
Vaccaro et al., 2008 USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP7 n=207 (7 mg/pt) ICBG n=86	single-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	rhBMP7 2.4 ICBG 2.7 (p=0.006)	rhBMP7 309 ICBG 471 (p=0.00004)	NSD but data not provided (p=0.529)	Significantly shorter OR time and less blood loss on average in rhBMP7 pts compared to ICBG
Vaccaro et al., 2008 USA **Lumbar Spine** Note: Long-term F/U study that includes all pts from Vaccaro et al., 2004, and Vaccaro et al., 2005	Multicenter, nonblinded RCT	rhBMP7 n=24 (7 mg/pt) ICBG n=12	single-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	rhBMP7 2.3±0.7 (0.8-3.7) ICBG 2.6±0.5) (1.9-3.6) (Data from Vaccaro et al., 2005, rec# 7310)	NR	rhBMP7 3.9±1.7 (2-10) ICBG 4.3±2.0 (3-9) (Data from Vaccaro et al., 2005, rec# 7310)	No significant differences between pt groups
Baskin et al., 2003 USA **Cervical Spine**	Multicenter, nonblinded RCT	rhBMP2/ALG n=18 (0.6-1.2 mg/pt) ICBG/ALG n=15	single- or two-level cervical DDD	single- or two-level primary instrumented ACDF with rhBMP2/ALG or ICBG/ALG	rhBMP2/ALG 1.8 ICBG/ALG 1.8	rhBMP2/ALG 91 ICBG/ALG 123	rhBMP2/ALG 1.4 ICBG/ALG 1.1	No significant intergroup differences reported
Butterman	Prospective	rhBMP2/CRA	single- or	single- or	rhBMP2/CRA	rhBMP2/CRA	rhBMP2/CRA	Cervical

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Mean OR time (hr)	Mean estimated blood loss (mL)	Mean hospital LOS (days)	Comment
et al., 2008 USA **Cervical Spine**	nonrandomized cohorts of consecutive patients	n=30 (0.9-3.7 mg/pt) ICBG n=36	multiple-level cervical DDD	multi-level primary instrumented or uninstrumented ACDF with rhBMP2/CRA or ICBG	1.9±0.4 ICBG 1.9±0.4	65±51 ICBG 65±84	1.3±0.5 ICBG 1.2±0.4	swelling caused dysphagia that was more severe in rhBMP2/CRA group than ICBG group, at 4 days after surgery and persisting for 21 days
Crawford et al., 2009 USA **Cervical Spine**	Retrospective cohort of consecutive patients	rhBMP2/BGE n=41 (4.2-12 mg/pt) ICBG n=36	single- or multi-level posterior cervical stenosis, ACDF nonunion, or unstable spondylosis	single- or multi-level instrumented posterior cervical spinal fusion with rhBMP2/BGE or ICBG	rhBMP2/BGE 2.8±1.0 ICBG 2.7±0.9	rhBMP2/BGE 275±224 ICBG 337±317	rhBMP2/BGE 4.2±2.6 ICBG 3.5±1.2	No significant differences reported between groups
Smucker et al., 2006 USA **Cervical Spine**	Retrospective case-control	rhBMP2/CRA n=69 (dose NR)	NR	single- or multi-level instrumented ACDF with rhBMP2/CRA or CRA alone	NR	NR	NR	Bivariate unadjusted logistic regression model showed significant association between cervical

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Mean OR time (hr)	Mean estimated blood loss (mL)	Mean hospital LOS (days)	Comment
								swelling and rhBMP2 ($p < 0.0001$), C4-C5 level surgery ($p=0.003$), age ≥ 50 years ($p=0.003$), surgery at ≥ 3 levels ($p=0.007$), combined sugery ($p=0.04$)
								Adjustment for demographic differences showed only rhBMP2 use was significantly associated with cervical swelling (OR 10.1, 95% CI 3.4, 29.7, $p < 0.0001$)
								Timing and presentation of cervical swelling in

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Mean OR time (hr)	Mean estimated blood loss (mL)	Mean hospital LOS (days)	Comment
		CRA n=165						rhBMP2 recipients was reported distinct from that typically seen after ACDF, usually about 4 days after surgery and qualitatively different
Vaidya et al., 2007 USA **Cervical Spine**	Retrospective cohort of consecutive patients	rhBMP2 n=22 (1-3 mg/pt) ALG/DBM n=24	single- or multiple-level cervical DDD with radiculopathy or myelopathy	single- or multi-level primary instrumented ACDF with interbody fusion cages rhBMP2 on ACS or ALG/DBM	NR	NR	rhBMP2 2.9 (1-9) ALG/DBM 2.3 (1-6)	Cervical swelling was significantly greater in the rhBMP2 group compared to the ALG/DBM group for 6 weeks postsurgery

Appendix Table 6. Comparative studies reported in the AHRQ HTA evaluating BMPs in spinal fusion: radiographic outcomes

Note. Abstraction tables copied directly from the AHRQ HTA report except that the references were changed to correspond to the current report. In addition, adverse events and complications were omitted as they were reported elsewhere.

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
On-label use								
Boden et al., 2000 USA **Lumbar spine**	Multicenter, nonblinded RCT	rhBMP2 (4.2-8.4 mg/pt) n=11 ICBG n=3	single-level DDD	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	3, 6, 12, 24 mos. rhBMP2 91, 100, 100, 100 ICBG 67 at all times	NR	Plain radiograph: < 5 degrees of angular motion on flexion-extension film, and absence of radiolucent lines covering 50% or more of implant surfaces CT: presence of continuous trabecular bone growing through both cages Fusion success	No evidence of clinically significant (1 mm) graft subsidence in either group, no anteroposteri or migration or rotation

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
							required agreement among all 5 independent readers unaware of treatment	
Burkus et al., 2002 USA **Lumbar spine**	Multicenter, nonblinded RCT	rhBMP2 (4.2-8.4 mg/pt) n=143 ICBG n=136	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	6, 12, 24 mos rhBMP2 97, 97, 94 ICBG 96, 93, 89	NR	Plain radiograph: < 3mm translation, < 5 degrees angular motion on flexion-extension film, and absence of radiolucent lines covering 50% or more of implant surfaces CT: presence of continuous trabecular bone growing through both cages	Secondary surgeries were classified as fusion failures regardless of independent radiologic assessment

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
							Fusion evaluated by two independent radiologists who were unaware of treatment, a third was consulted for adjudication of disagreement	
Burkus et al., 2003 (Integrated analysis) **Lumbar spine** Note: may include pts in Burkus et al., 2003 ("Radio-graphic assessment …")	Retrospective combined comparative analysis	rhBMP2 n=277 (dose NR) ICBG n=402	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages	6, 12, 24 mos rhBMP2 95, 96, 94 ICBG 96, 93, 89 (p=0.022 at 24 mos)	NR	Same as Burkus et al., 2002 (rec#11620)	Fusion success difference at 24 mos. statistically significant by ANCOVA
Off-label use								
Boden et al., (2002)	Multicenter, nonblinded	rhBMP2/CR M	single-level lumbar DDD	single-level primary	24 mos. (22/27 pts)	NR	Presence of bridging	By 12 mos. and

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
USA **Lumbar Spine**	RCT	plus Texas Scottish Rite Hospital (TSRH) Spinal System (TSRHSS) n=11		instrumented posterolateral lumbar fusion plus rhBMP2 ICBG	rhBMP2/CRM /TSRHSS 100		trabecular bone between the transverse processes, absence of motion, defined as 3 mm or less of translation and < 5 degrees of angular motion on flexion-extension views, and absence of radiolucent lines through the fusion mass Fusion evaluated by two independent radiologists who were unaware of treatment	continuing at 24 mos, the opacity of the ceramic CRM changed from a pale gray speckled pattern to a more uniform, well-marginated whiter mass
		(40 mg/pt) rhBMP2/CRM alone n=11			rhBMP2/CRM alone 100			
		(40 mg/pt) ICBG plus TSRHSS n=5			ICBG/TSRHSS 40 (p=0.018, 0.028 in BMP2 groups vs ICBG)			

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
Burkus et al., (2005) USA **Lumbar Spine** Note: includes all pts from Burkus et al., 2002, rec# 11510; same pts as Burkus et al., 2006, rec# 6640	Multicenter, nonblinded RCT	rhBMP2 n=79 (8-12 mg/pt) ICBG N=52	single-level lumbar lumbar DDD	primary single-level anterior lumbar fusion with a pair of threaded allograft cortical bone dowels (CBD) plus rhBMP2 or ICBG	6, 12, 24 mos rhBMP2 96, 99, 98 ICBG 85, 89, 76 (p=0.047, 0.035, < 0.001)	NR	Presence of bridging bone connecting adjacent vertebral bodies, either through the FRA or around the FRA, < 5 degrees of angular motion, ≤ 3 mm translation, and absence of radiolucent lines around > 50% of the graft Fusion evaluated by two independent radiologists who were unaware of treatment, a third was consulted for	Fusion was deemed successful only if all criteria were met In the ICBG group, no patient had a fracture, migration, or extrusion of the FRA 14 (18%) of 79 patients in the rhBMP2 group had transient localized areas of bone remodeling in the vertebral body adjacent to a FRA, visible between 3 and 12 mos. postsurgery,

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
							adjudication of disagreement	but resolved by 24 mos
Dawson et al., 2009 USA **Lumbar spine**	Multicenter nonblinded RCT	rhBMP2/CRM n=25 (12 mg/pt) ICBG n=21	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	6, 12, 24 mos rhBMP2/CRM 91, 89, 95 ICBG 58, 65, 67 (p=0.032 at 6 mos)	NR	Presence of bridging trabecular bone between the transverse processes, absence of motion, defined as 3 mm or less of translation and < 5 degrees of angular motion on flexion-extension views, and absence of radiolucent lines through the fusion mass Fusion evaluated by two	Thin-cut CT showed progressive formation of bridging bone across the transverse processes and incorporation of the ceramic component

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
							independent radiologists who were unaware of treatment, a third was consulted for adjudication of disagreement	
Dimar et al., (2009) USA **Lumbar Spine** Note: contains pts in Glassman et al., 2007, rec# 4040; Dimar et al., 2006 rec#5480; Glassman et al., 2005, rec# 8040	Multicenter nonblinded RCT	rhBMP2/CRM n=239 (40 mg/pt) ICBG n=224	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	6, 12, 24 mos rhBMP2/CRM 79, 88, 96 ICBG 65, 83, 89 (p=0.002, 0.107, 0.014)	NR	Presence of bridging trabecular bone between the transverse processes, absence of motion, defined as 3 mm or less of translation and < 5 degrees of angular motion on flexion-extension views, and absence of radiolucent	Thin-cut CT showed progressive formation of bridging bone across the transverse processes

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
							lines through the fusion mass Fusion evaluated by two independent radiologists who were unaware of treatment, a third was consulted for adjudication of disagreement	
Glassman et al., (2008) USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP2/ACS n=50 (dose not reported) ICBG n=52	single- or multi-level lumbar DDD	single- or multi-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	rhBMP2 86 Average CT fusion grade at 24 mos rhBMP2 4.3±1.3 ICBG 71 Average CT fusion grade at 24 mos ICBG 3.8±0.9	NR	CT fusion rating scale: Grade 1=no fusion Grade 2=partial or limited unilateral fusion Grade 3=partial or limited bilateral fusion	Fusion grade a composite score from 3 reviewers of CT scans

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
					(p=0.030)		Grade 4=solid unilateral fusion Grade 5=solid bilateral fusion Fusion evaluated independently by 3 orthopedic spine surgeons unaware of treatment	
Haid et al., (2004) USA **Lumbar Spine**	Multicenter, nonblinded RCT	rhBMP2 n=34 (4.2-8.4) ICBG N=33	single-level lumbar DDD	single-level primary posterior lumbar interbody fusion (PLIF) interbody fusion cages plus rhBMP2 or ICBG	6, 12, 24 mos rhBMP2 93, 85, 92 ICBG 93, 92, 78	NR	Presence of bridging bone connecting adjacent vertebral bodies, < 5 degrees of angular motion, ≤ 3 mm translation, and absence of radiolucent lines around > 50% of the	Secondary surgeries were classified as fusion failures regardless of independent radiologic assessment New bone formation extending outside the

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
							graft Fusion evaluated by two independent radiologists who were unaware of treatment, a third was consulted for adjudication of disagreement	disc space and into the spinal canal or neuroforamin a was observed in 24 rhBMP2 (71) and 4 (12) ICBG recipients (p < 0.0001) but was not correlated with recurrence or increase in leg pain from the preoperative status
Glassman et al., (2007) USA **Lumbar Spine**	Retrospective with historical control group	rhBMP2 n=91 (12 mg/pt) ICBG n=35	single- and multi-level lumbar DDD, degenerative scoliosis, postdiscecto my instability, spinal stenosis, adjacent level	single- or multi-level primary or revision instrumented posterolateral lumbar fusion	rhBMP2 24 mos 87 of 91 (96) ICBG 24 mos 30 of 35 (86)	NR	Plain radiographs: fusion mass graded as solid fusion, probabale fusion, or nonunion CT fusion	Fusion grade a composite score from 2 reviewers of CT scans **Outcomes corrected by Spectrum.**

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
			degeneration				rating scale: Grade 1=no fusion Grade 2=partial or limited unilateral fusion Grade 3=partial or limited bilateral fusion Grade 4=solid unilateral fusion Grade 5=solid bilateral fusion Fusion evaluated by two independent radiologists who were unaware of treatment	
Mumma-neni et al., 2004	Retrospective single-center cohort study	rhBMP2/AGB n=25 (8.4 mg/pt)	single- or multi-level lumbar DDD	single- or multi-level primary	rhBMP2/AGB 96 at average 8 mos. F/U	rhBMP2/AGB 3.6±2.0 (1-9)	Presence of bridging bone connecting	Only used plain radiographs

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
USA **Lumbar Spine**		ICBG N=19		transforaminal lumbar interbody fusion (TLIF) with interbody fusion cages with rhBMP2 plus AGB or ICBG alone	ICBG 95 at average 11 mos. F/U	ICBG 6.4±2.4 (3-12)	adjacent vertebral bodies, lack of motion on dynamic flexion-extension radiographs, absence of halo around screws Fusion analysis method not mentioned	for fusion studies
Pradhan et al., 2006 USA **Lumbar Spine**	Prospective consecutive patient single-center cohort study	rhBMP2 n=9 (dose NR) ICBG n=27	single-level lumbar DDD	single-level primary anterior lumbar interbody fusion (ALIF) with femoral ring allograft (FRA) plus rhBMP2 or ICBG	24 mos rhBMP2 4 of 9 (44) Non-unions rhBMP 5 (56) 24 mos ICBG 17 of 27 (63) Non-unions ICBG 10 (37)	NR	Presence of bridging bone connecting adjacent vertebral bodies, either through the FRA or around the FRA, < 5 degrees of angular motion, ≤ 3 mm translation,	Fusion was deemed successful only if all criteria were met Graft and endplate resorption reported to occur earlier and more aggressively in pts treated

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
							and absence of radiolucent lines around > 50% of the graft Fusion evaluated by a radiologist who was unaware of treatment	with rhBMP2 compared with ICBG, which may be related to number of non-unions and delayed unions
Singh et al., 2006 USA **Lumbar Spine**	Prospective single-center case-matched cohort study	rhBMP2/ICBG n=39 (12-36 mg/pt) ICBG N=11	single- or multi-level lumbar DDD	single- or multi-level primary instrumented posterolateral lumbar fusion with rhBMP2 plus ICBG or ICBG alone	24 mos rhBMP2/ICBG 94 (68 of 70 levels) ICBG 77 (17 of 22 levels) ($p < 0.05$)	NR	Presence of continuous trabecular bone between intertransverse processes, cortication at the peripheral edge of the fusion mass, and absence of identifiable radiographic cleft on CT assessment Fusion evaluated by	Fusion qualitry was subjectively assessed as excellent in 92% of rhBMP2/ICBG recipients and 27% of ICBG recipients ($p < 0.05$)

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
							two orthopedic surgeons and a radiologist, all unaware of treatment	
Slosar et al., 2007 USA **Lumbar Spine**	Prospective consecutive patient single-center cohort study	rhBMP2 n=45 (3-9 mg/pt) ALG N=30	single- or multi-level lumbar DDD	single- or multi-level primary instrumented anterior lumbar interbody fusion (ALIF) with femoral ring allograft (FRA) plus rhBMP2 or allograft bone chips (ALG)	6, 12, 24 mos rhBMP2 79, 96, 99 ALG 23, 73, 82 (p < 0.001 at all times)	NR	Molinari-Bridwell grading (Molinari et al., 1999) scale used: Grade 1: fused with remodeling and trabeculae present Grade 2: Graft intact, not fully remodeled and incorporated, no lucency Grade 3: Graft intact, potential lucency present at top or bottom of	No osteolysis or fragmentations of FRA were observed

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
							graft Grade 4: Fusion absent with collapse/resorption of graft Grades 1-2 were considered fused, Grades 3-4 considered not fused All studies were reviewed by independent reviewers unaware of treatment	
Johnsson et al., 2002 Sweden **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP7 n=10 (7 mg/pt)	single-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	Radiographic fusion 12 mos rhBMP7 60 bilateral bridging bone 30 partial bone formation	NR	Bone formation classified as radiographic evidence of bilaterally bridging bone, partial bone	RSA analysis showed no significant differences in L5 stabilization or movement

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
					10 no bone formation		formation, or no bone formation	
		ICBG n=10			ICBG 80 bilateral bridging bone			
					20 partial bone formation			
Kanayama et al., 2006 Japan, USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP7 n=9 (7 mg/pt) AGB/CRM n=10	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion with rhBMP7 or AGB/CRM	Radiographic fusion criteria at 15.3 mos rhBMP7 78	NR	Presence of bridging bone on CT scan in posterolateral lumbar area, ≤ 5 degrees of angulation and ≤ 2 mm of translation at the index level	No significant differences in fusion,but small pt numbers limit ersults
					Surgical evidence of solid fusion rhBMP7 57 (4 of 7)			
		AGB/CRM n=10			Radiographic fusion criteria at 15.3 mos AGB/CRM 90			
					Surgical evidence of solid fusion AGB/CRM 78 (7 of 9)			
Vaccaro et al., 2008 USA	Multicenter nonblinded RCT	rhBMP7 n=207 (7 mg/pt)	single-level lumbar DDD	single-level primary uninstrument	Bridging bone (CT) 36+ mos	NR	Presence of new bone formation	Overall radiographic comprised 3

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
Lumbar Spine				ed posterolateral lumbar fusion with rhBMP7 or ICBG	rhBMP2 75 ≤ 5 degrees angulation (plain film) rhBMP7 69 ≤ 3 mm translation (plain film) rhBMP7 76		bridging across the transverse processes, angulation ≤ 5 degrees, and ≤ 3 mm of translation were required Fusion evaluated independently by 2 primary spine surgeons unaware of treatment, a third was consulted for adjudication of disagreement	components necessary to define fusion No significant differences seen in fusion parameters at 36+ mos. F/U
		ICBG n=86			Bridging bone (CT) 36+ mos ICBG 77 ≤ 5 degrees angulation (plain film) ICBG 68 ≤ 3 mm translation (plain film) ICBG 75			
Vaccaro et al., 2008 USA	Multicenter, nonblinded RCT	rhBMP7 n=24 (7 mg/pt)	single-level lumbar DDD	single-level primary uninstrumented	Solid fusion 48 mos rhBMP7 69 (11 of 16	NR	Complete bridging bone between transverse	Both groups showed equivalent reductions in

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
Lumbar Spine Note: Long-term F/U study that includes all pts from Vaccaro et al., 2004, and Vaccaro et al., 2005		ICBG n=12		posterolateral lumbar fusion with rhBMP7 or ICBG	with data) Bridging bone 48 mos rhBMP7 81 (13 of 16 with data) Solid fusion ICBG 50 (3 of 6 with data) Bridging bone 48 mos ICBG 50 (3 of 6 with data)		processes, ≤ 5 degrees of angulation and ≤ 2 mm of translation Fusion evaluated independently by 2 neuroradiologists unaware of treatment, a third was consulted for adjudication of disagreement	disc height as well as angular and translational motion at the treated level
Baskin et al., 2003 USA **Cervical Spine**	Multicenter, nonblinded RCT	rhBMP2/ALG n=18 (0.6-1.2 mg/pt) ICBG/ALG n=15	single- or two-level cervical DDD	single- or two-level primary instrumented ACDF with rhBMP2/ALG or ICBG/ALG	6, 12, 24 mos rhBMP2/ALG 100 at all times ICBG/ALG 100 at all times	NR	Plain radiograph: < 4 degrees difference in angular motion between flexion and extension, no radiolucency > 2 mm thick covering > 50% of the	Two pts in rhBMP2/ALG and one in the ICBG/ALG group demonstrated bone formation immediately anterior to segments adjacent to

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
							inferior or superior graft surface, presence of bridging trabecular bone CT: presence of bridging trabecular bone	the index level
Butterman et al., 2008 USA **Cervical Spine**	Prospective nonrandomized cohorts of consecutive patients	rhBMP2/CRA n=30 (0.9-3.7 mg/pt) ICBG n=36	single- or multiple-level cervical DDD	single- or multi-level primary instrumented or uninstrumented ACDF with rhBMP2/CRA or ICBG	NR	NR	Plain films: Presence of bridging trabecular bone across disc space, < 1 mm gapping of spinous processes on flexion-extension films and selected high-resolution CT scans	2 pseudarthroses in ICBG group, 1 in the rhBMP2/CRA group
Crawford et al., 2009 USA	Retrospective cohort of consecutive patients	rhBMP2/BGE n=41 (4.2-12 mg/pt)	single- or multi-level posterior cervical	single- or multi-level instrumented posterior	NR	NR	NR	

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
Cervical Spine		ICBG n=36	stenosis, ACDF nonunion, or unstable spondylosis	cervical spinal fusion with rhBMP2/BGE or ICBG				
Smucker et al., 2006 USA **Cervical Spine**	Retrospective case-control	rhBMP2/CRA n=69 (dose NR) CRA n=165	NR	single- or multi-level instrumented ACDF with rhBMP2/CRA or CRA alone	NR	NR	NR	
Vaidya et al., 2007 USA **Cervical Spine**	Retrospective cohort of consecutive patients	rhBMP2 n=22 (1-3 mg/pt) ALG/DBM n=24	single- or multiple-level cervical DDD with radiculopathy or myelopathy	single- or multi-level primary instrumented ACDF with interbody fusion cages rhBMP2 on ACS or ALG/DBM	rhBMP2 100 ALG/DBM 96	NR	For the rhBMP2 group, bone formation was assessed as no new bone, visible new bone, possible fusion, and probable fusion For the ALG/DBM group fusion was assessed at the graft	End plate resorption was noted in 100% of the levels where rhBMP2 was used, starting at 1.5 mos. and lasting until 6 mos

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Successful outcome (%) (p-value)	Time to successful outcome mn ± SD (rng) (p-value)	Definition of successful outcome	Comment
							endplate junction, classified as not united, possibly united, and probably united	

Appendix Table 7. Comparative studies reported in the AHRQ HTA evaluating BMPs in spinal fusion: pain outcomes

Note. Abstraction tables copied directly from the AHRQ HTA report except that the references were changed to correspond to the current report. In addition, adverse events and complications were omitted as they were reported elsewhere.

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
On-label use							
Boden et al., 2000 USA **Lumbar spine**	Multicenter, nonblinded RCT	rhBMP2 (4.2-8.4 mg/pt) n=11	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	Oswestry DI Mean score improvement (points) 3, 6, 12, 24 mos rhBMP2 9, 12, 22, 25	Oswestry DI ≥ 15% improvement 3, 6, 12, 24 mos rhBMP2 55, 64, 91, 91	Success for ODI defined as ≥ 15% improvement over baseline score
		ICBG n=3			Oswestry DI Mean score improvement (points) 3, 6, 12, 24 mos ICBG 35, -18, 7, 8, 15	ICBG 0, 67, 67, 67	
					Iliac crest pain postharvest NR		
Burkus et al., 2002 USA **Lumbar spine**	Multicenter, nonblinded RCT	rhBMP2 (4.2-8.4 mg/pt) n=143	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	Oswestry DI Mean score improvement (points) 1.5, 3, 6, 12, 24 mos rhBMP2 12, 20, 25, 28, 30	Oswestry DI 12, 24 mos rhBMP2 85, 84	Success for ODI defined as ≥ 15% improvement over baseline score Both groups showed significant improvements from baseline, but there were no significant
					Back pain Mean score improvement (points) 1.5, 3, 6, 12, 24 mos rhBMP2 6.5, 7.1, 7.2, 7.8, 8.5	Back pain (> 3 point improvement) 1.5, 3, 6, 12, 24 mos rhBMP 77, 74, 78, 79, 75	
					Leg pain Mean score improvement (points) 1.5, 3, 6, 12, 24 mos rhBMP2	Leg pain (> 3 point improvement if baseline score > 10 points, or maintenance of	

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
					5.0, 5.7, 6.2, 6.2, 6.2	score if < 10) 12, 24 mos rhBMP2 72, 80	differences between groups in mean score or rates
		ICBG n=136			Oswestry DI Mean score improvement (points) 1.5, 3, 6, 12, 24 mos ICBG 55, 14, 21, 26, 29, 31	Oswestry DI 12, 24 mos ICBG 86, 82	
					Back pain Mean score improvement (points) 1.5, 3, 6, 12, 24 mos ICBG 7.3, 7.1, 7.2, 7.7, 8.2	Back pain (> 3 point improvement) 1.5, 3, 6, 12, 24 mos ICBG 76, 78, 72, 73, 79	
					Leg pain Mean score improvement (points) 1.5, 3, 6, 12, 24 mos ICBG 4.1, 5.7, 6.2, 5.9, 6.2	Leg pain (> 3 point improvement if baseline score > 10 points, or maintenance of score if < 10) 12, 24 mos ICBG 73, 74	
					Iliac crest pain postharvest Mean score (20 point VAS) 0, 24 mos 12.7, 1.8	Iliac crest pain postharvest % at 24 mos 32	
Burkus et al., 2003 (Integrated analysis) **Lumbar spine** Note: may include pts in Burkus et al.,	Retrospective combined comparative analysis	rhBMP2 n=277 (dose NR)	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages	Oswestry DI Mean score improvement (points) 3, 6, 12, 24 mos rhBMP2 31, 26, 30, 31	NR	Both groups improved over time
					SF-36 pain index subscale Mean score improvement (points)		

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
2003 ("Radio-graphic assessment…")					3, 6, 12, 24 mos rhBMP2 27, 32, 36, 39		
		ICBG N=402			Oswestry DI Mean score improvement (points) 3, 6, 12, 24 mos ICBG 5, 20, 23, 26 (p=0.0041, 0.0053, 0.0013, 0.0023 rhBMP2 vs ICBG)		
					SF-36 pain index subscale Mean score improvement (points) 3, 6, 12, 24 mos ICBG 20, 24, 29, 33 (p=0.0002 at 3, 6, 12 mos. and 0.0008 at 24 mos, rhBMP2 vs ICBG)		
					Iliac crest pain postharvest NR		
Off-label use							
Boden et al., (2002) USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP2/CRM plus Texas Scottish Rite Hospital (TSRH) Spinal System (TSRHSS) n=11	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 ICBG	Oswestry DI Mean score improvement (points) 1.5, 3, 6, 17 mos rhBMP2/CRM/TSRHSS ~3, ~18, ~20, ~13	Oswestry DI ≥ 15% improvement 1.5, 3, 6, 17 mos rhBMP2/CRM/TSRHSS ~38, ~80, ~80, ~65	All pain outcomes showed significant improvement in both groups at 17-24 mos. but no significant intergroup differences except for SF-36 score at 17 mos
					Back pain Mean score improvement (points) 1.5, 3, 6, 17 mos rhBMP2/CRM/TSRHSS ~6, ~8, ~7, ~5		
					Leg pain Mean score improvement (points) 1.5, 3, 6, 17 mos		

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
					rhBMP2/CRM/TSRHSS ~3, ~4, ~1, ~3		
					SF-36 bodily pain subscale Mean score improvement (points) 1.5, 3, 6, 17 mos rhBMP2/CRM/TSRHSS ~3, ~10, ~23, ~15		
		(40 mg/pt) rhBMP2/CRM alone n=11			Oswestry DI Mean score improvement (points) 1.5, 3, 6, 17 mos rhBMP2/CRM alone ~19, ~22, ~25, ~29	rhBMP2 alone ~88, ~88, ~88, ~100	
					Back pain Mean score improvement (points) 1.5, 3, 6, 17 mos rhBMP2/CRM alone ~8, ~9, ~9, ~10		
					Leg pain Mean score improvement (points) 1.5, 3, 6, 17 mos rhBMP2/CRM ~8, ~9, ~7, ~9		
					SF-36 bodily pain subscale Mean score improvement (points) 1.5, 3, 6, 17 mos rhBMP2/CRM alone ~22, ~32, ~35, ~35		
		(40 mg/pt) ICBG plus TSRHSS n=5			Oswestry DI Mean score improvement (points) 1.5, 3, 6, 17 mos ICBG/TSRHSS	ICBG/TSRHSS ~80, ~60, ~80, ~80	

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
					~10, ~15, ~17, ~25 Back pain Mean score improvement (points) 1.5, 3, 6, 17 mos ICBG/TSRHSS ~7, ~5, ~4, ~5 Leg pain Mean score improvement (points) 1.5, 3, 6, 17 mos rhBMP2/CRM/TSRHSS ICBG/TSRHSS ~7, ~3, ~3, ~4 SF-36 bodily pain subscale Mean score improvement (points) 1.5, 3, 6, 17 mos ICBG/TSRHSS ~3, ~10, ~23, ~15 (rhBMP2/CRM alone, p=0.049 vs the other 2 groups)		
Burkus et al., (2005) USA **Lumbar Spine** Note: includes all pts from Burkus et al., 2002, rec# 11510; same pts as Burkus et al., 2006, rec# 6640	Multicenter, nonblinded RCT	rhBMP2 n=79 (8-12 mg/pt)	single-level lumbar DDD	primary single-level anterior lumbar fusion with a pair of threaded allograft cortical bone dowels (CBD) plus rhBMP2 or ICBG	Oswestry DI Mean score improvement (points) 6, 12, 24 mos rhBMP2 32.4, 33.0, 33.4 Back pain Mean score improvement (points) 6, 12, 24 mos rhBMP2 9.2, 9.2, 8.6 Leg pain Mean score improvement (points) 6, 12, 24 mos rhBMP2 7.7, 7.5, 6.8	NR	**NOTE: all data added to this chart by Spectrum (none supplied by AHRQ)** Both groups had statistically significant improvement in the mean ODI, back, and leg pain scores

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
		ICBG N=52			Oswestry DI Mean score improvement (points) 6, 12, 24 mos ICBG 25.8, 27.0, 27.0 *P* = .031, .074, .119 Back pain Mean score improvement (points) 6, 12, 24 mos ICBG 7.7, 7.3, 7.1 *P* = .006, .007, .032 Leg pain Mean score improvement (points) 6, 12, 24 mos ICBG 7.3, 6.2, 4.9 *P* = .043, .011, .011		compared to preoperative values Statistically signficant intergroup differences favoring rhBMP2 seen in all three indexes at specific times
Dawson et al., 2009 USA **Lumbar spine**	Multicenter nonblinded RCT	rhBMP2/CRM n=25 (12 mg/pt)	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	Oswestry DI Mean score improvement (points) 24 mos rhBMP2/CRM 28 Back pain Mean score improvement (points) 24 mos rhBMP2/CRM 9.6 Leg pain Mean score improvement (points) 24 mos rhBMP2/CRM 9.3	Oswestry DI > 20% improvement 24 mos rhBMP2/CRM 91	Overall success rate was 81% in rhBMP2/CRM group and 55% in the ICBG group (p NSD)

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
		ICBG n=21			Oswestry DI Mean score improvement (points) 24 mos ICBG 23 (*P* = .953) Back pain Mean score improvement (points) 24 mos ICBG 7.2 Leg pain Mean score improvement (points) 24 mos ICBG 7.2 Iliac crest pain postharvest NR	ICBG 70 (*P* = .532)	
Dimar et al., (2009) USA **Lumbar Spine** Note: contains pts in Glassman et al., 2007, rec# 4040; Dimar et al., 2006 rec# 5480; Glassman et al., 2005, rec# 8040	Multicenter nonblinded RCT	rhBMP2/CRM n=239 (40 mg/pt)	single- or multi-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	Oswestry DI Mean score improvement (points) 24 mos rhBMP2 estimated from graph 27 Back pain Mean score improvement (points) 24 mos rhBMP2 estimated from graph 9 Leg pain Mean score improvement (points) 24 mos rhBMP2 estimated from graph 8	NR	All pain outcomes (ODI, back pain, leg pain) showed significant improvement in both groups at 24 mos. but no significant intergroup differences **NOTE: all data added to this chart by Spectrum (none supplied**
		ICBG n=224			Oswestry DI Mean score improvement (points)		

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
					24 mos estimated from graph 26		**by AHRQ)**
					Back pain Mean score improvement (points) 24 mos rhBMP2 estimated from graph 8		
					Leg pain Mean score improvement (points) 24 mos rhBMP2 estimated from graph 8		
Glassman et al., (2008) USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP2/ACS n=50 (dose not reported)	single-level lumbar DDD	single- or multi-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	Oswestry DI Mean score improvement (points) 3, 6, 12, 24 mos rhBMP2 14, 18, 19, 15	NR	Mean pain scores were similar in both groups at all time intervals, with statistically significant improvement compared to preoperative mean scores but no significant intergroup differences
					Back pain Mean score improvement (points) 1.5, 6, 12, 24 rhBMP2 4.3, 4.1, 4.1, 3.1		
					Leg pain Mean score improvement (points) 1.5, 6, 12, 24 mos rhBMP2 4.6, 4.4, 3.8, 3.6		
		ICBG n=52			Oswestry DI Mean score improvement (points) 3, 6, 12, 24 mos ICBG 13, 17, 18, 13		
					Back pain		

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
					Mean score improvement (points) 1.5, 6, 12, 24 ICBG 4.0, 4.0. 3.9, 3.0		
					Leg pain Mean score improvement (points) 1.5, 6, 12, 24 mos ICBG 4.1, 4.2, 3.9, 3.1		
					Iliac crest pain postharvest NR		
Haid et al., (2004) USA **Lumbar Spine**	Multicenter, nonblinded RCT	rhBMP2 n=34 (4.2-8.4)	single- or multi-level lumbar DDD	single-level primary posterior lumbar interbody fusion (PLIF) interbody fusion cages plus rhBMP2 or ICBG	Oswestry DI Mean score improvement (points) 24 mos rhBMP2 30	Oswestry DI ≥ 15% improvement 24 mos rhBMP2 69	Both groups had statistically significant improvements in mean ODI, back, and leg pain at all times compared to preoperative values
					Back pain Mean score improvement (points) 24 mos rhBMP2 9		
					Leg pain Mean score improvement (points) 24 mos rhBMP2 7.7		
		ICBG N=33			Oswestry DI Mean score improvement (points) 24 mos ICBG 25	ICBG 56	
					Back pain Mean score improvement (points)		

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
					24 mos ICBG 4.5 (p=0.009)		
					Leg pain Mean score improvement (points) 24 mos ICBG 6.5		
					Iliac crest pain postharvest Mean score (points) 24 mos 5.5		
					% with pain at 24 mos 60		
Glassman et al., (2007) USA **Lumbar Spine**	Retrospective with historical control group	rhBMP2 n=91 (12 mg/pt) ICBG n=35	single-level lumbar DDD	single- or multi-level primary or revision instrumented posterolateral lumbar fusion	NR	NR	Study only reported fusion data
Mummaneni et al., 2004 USA **Lumbar Spine**	Retrospective single-center cohort study	rhBMP2/AGB n=25 (8.4 mg/pt) ICBG N=19	single-level lumbar DDD	single- or multi-level primary transforaminal lumbar interbody fusion (TLIF) with interbody fusion cages with rhBMP2 plus AGB or ICBG alone	Prolo Scale Pain subscale Mean score at F/U (points) rhBMP2/AGB 3.8±0.9	NR	Statistical analysis not done
					Prolo Scale Pain subscale Mean score at F/U (points) ICBG 4.0±0.7		
					% with pain 6 mos		

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
					58 Mean pain score (points) 6 mos 5		
Pradhan et al., 2006 USA **Lumbar Spine**	Prospective consecutive patient single-center cohort study	rhBMP2 n=9 (dose NR) ICBG n=27	single- and multi-level lumbar DDD, degenerative scoliosis, postdiscectomy instability, spinal stenosis, adjacent level degeneration	single-level primary anterior lumbar interbody fusion (ALIF) with femoral ring allograft (FRA) plus rhBMP2 or ICBG	NR Iliac crest pain NR	NR	Study only reported fusion data
Singh et al., 2006 USA **Lumbar Spine**	Prospective single-center case-matched cohort study	rhBMP2/ICBG n=39 (12-36 mg/pt) ICBG N=11	single- or multi-level lumbar DDD	single- or multi-level primary instrumented posterolateral lumbar fusion with rhBMP2 plus ICBG or ICBG alone	NR Iliac crest pain NR	NR	
Slosar et al., 2007 USA **Lumbar Spine**	Prospective consecutive patient single-center cohort study	rhBMP2 n=45 (3-9 mg/pt)	single-level lumbar lumbar DDD	single- or multi-level primary instrumented anterior lumbar interbody fusion (ALIF) with femoral ring allograft (FRA) plus	Oswestry DI Mean score improvement (points) 6, 12, 24 mos rhBMP2 27, 30, 33 NRS (undefined) Mean score improvement (points) 6, 12, 24 mos rhBMP2 4.2, 4.7, 4.8	NR	Both groups had statistically significant improvements in mean ODI and NRS at all times compared to preoperative values

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
		ALG N=30		rhBMP2 or allograft bone chips (ALG)	Oswestry DI Mean score improvement (points) 6, 12, 24 mos ALG 17, 26, 30 (p < 0.001 at 6 mos) NRS (undefined) Mean score improvement (points) 6, 12, 24 mos ALG 2.8, 4.4, 4.3 (p < 0.001 at 6 mos)		
Johnsson et al., 2002 Sweden **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP7 n=10 (7 mg/pt) ICBG n=10	single-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	NR Iliac crest pain	Subjective evaluation of back pain 12 mos rhBMP7 None (4 pts) Minor w/out medication (4 pts) Major with medication (2) Subjective evaluation of back pain 12 mos ICBG None (5 pts) Minor w/out medication (2 pts) Major with medication (3 pts)	Patients had similar pain outcomes, but no statistical analysis was done
Kanayama et al., 2006 Japan, USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP7 n=9 (7 mg/pt)	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion	Oswestry DI Mean score improvement (points) 3, 6, 9, 12 mos rhBMP7 ~15, ~23, ~16, ~17	NR	Both groups had signficant decreases in pain from baseline

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
		AGB/CRM n=10		with rhBMP7 or AGB/CRM	AGB/CRM ~17, ~31, ~24, ~24		(p < 0.05, ANOVA), but NSD between groups
Vaccaro et al., 2008 USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP7 n=207 (7 mg/pt)	single-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	Oswestry DI mean percent improvement from baseline 36+ mos rhBMP7 52	Modified Overall Success 36+ mos rhBMP7 47	Both groups had significant decreases in pain from baseline levels
					VAS scores 36+ mos NSD	Oswestry DI ≥ 20% improvement 36+ mos rhBMP7 69	
					SF-36 scores NSD		
		ICBG n=86			Oswestry DI mean percent improvement from baseline 36+ mos ICBG 54	Modified Overall Success 36+ mos ICBG 47 (p for noninferiority=0.025)	
					Iliac crest pain postharvest % with pain 12, 24, 36+ mos 44, 45, 35	Oswestry DI ≥ 20% improvement 36+ mos ICBG 77	
					Mean pain score (points) 1.5, 12, 24, 36+ mos 2.1, 1.6, 1.2, 1.1		
Vaccaro et al., 2008 USA **Lumbar Spine** Note: Long-term F/U	Multicenter, nonblinded RCT	rhBMP7 n=24 (7 mg/pt)	single- or multi-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7	Oswestry DI mean score NR	Oswestry DI ≥ 20% improvement 48 mos rhBMP7 74 (14 of 19 with data) (95% CI, 49, 91)	Overall success is a composite measure comprising definitive spinal fusion, minimum 20%
						Overall success	

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
study that includes all pts from Vaccaro et al., 2004, and Vaccaro et al., 2005				or ICBG		48 mos rhBMP7 62 (10 of 16 with data)	improvement in Oswestry DI, and absence of surgical retreatment
						Overall success 48 mos, LOCF analysis rhBMP7 46 (95% CI, 26, 67)	
		ICBG n=12			Iliac crest pain NR	Oswestry DI ≥ 20% improvement 48 mos ICBG 57 (4 of 7 with data) (95% CI, 18, 90)	
						Overall success 48 mos ICBG 33 (2 of 6 with data)	
						Overall success 48 mos, LOCF analysis ICBG 25 (95% CI, 6-57)	
Baskin et al., 2003 USA **Cervical Spine**	Multicenter, nonblinded RCT	rhBMP2/ALG n=18 (0.6-1.2 mg/pt)	single- or two-level cervical DDD	single- or two-level primary instrumented ACDF with rhBMP2/ALG or ICBG/ALG	Neck Disability Index Mean score improvement (points) 1.5, 3, 6, 12, 24 mos rhBMP2/ALG 37, 39, 48, 46, 53	Neck pain 24 mos rhBMP2/ALG 100	Both groups showed significant improvements from baseline, but there were no significant differences between groups in mean score or rates
					Neck pain Mean score improvement (points) 1.5, 3, 6, 12, 24 mos rhBMP2/ALG 11, 11, 11, 12, 13		
					Arm pain		

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
					Mean score improvement (points) 1.5, 3, 6, 12, 24 mos rhBMP2/ALG 14, 14, 15, 14, 14		
		ICBG/ALG n=15			Neck Disability Index Mean score improvement (points) 1.5, 3, 6, 12, 24 mos ICBG/ALG 33, 34, 39, 41, 37 (p < 0.03 at 24 mos)	ICBG/ALG 100	
					Neck pain Mean score improvement (points) 1.5, 3, 6, 12, 24 mos ICBG/ALG 7, 8, 10, 9, 9		
					Arm pain Mean score improvement (points) 1.5, 3, 6, 12, 24 mos ICBG/ALG 9, 8, 10, 10, 8 (p < 0.03 at 24 mos)		
					Iliac crest pain postharvest 1.5, 6, 24mos Pain reported at each time, but not quantified		
Butterman et al., 2008 USA **Cervical Spine**	Prospective nonrandomized cohorts of consecutive patients	rhBMP2/CRA n=30 (0.9-3.7 mg/pt)	single- or multiple-level cervical DDD	single- or multi-level primary instrumented or uninstrumented ACDF with rhBMP2/CRA or ICBG	Oswestry Disability Index Mean score improvement (points) 7-12, 13-24, 25-36 mos rhBMP2/CRA ~14, ~25, ~30	NR	Both groups showed significant improvements from baseline, but there were no significant differences between groups
					Neck pain Mean score improvement (points) 7-12, 13-24, 25-36 mos rhBMP2/CRA ~4, ~4.5, ~5		
					Arm pain		

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
					Mean score improvement (points) 7-12, 13-24, 25-36 mos rhBMP2/CRA ~3.3, ~4.2, ~5.5		in mean score or rates
					Narcotic pain medication use (%) preop, 7-12, 13-24, 25-36 mos rhBMP2/CRA 53, 30, 23, 10		
		ICBG n=36			Oswestry Disability Index Mean score improvement (points) 7-12, 13-24, 25-36 mos ICBG ~11, ~17, ~31		
					Neck pain Mean score improvement (points) 7-12, 13-24, 25-36 mos ICBG ~4, ~4, ~5		
					Arm pain Mean score improvement (points) 7-12, 13-24, 25-36 mos ICBG ~3.9, ~3.8, ~4.8		
					Narcotic pain medication use (%) preop, 7-12, 13-24, 25-36 mos ICBG 61, 39, 19, 6		
					Iliac crest pain postharvest		
Crawford et al., 2009 USA **Cervical Spine**	Retrospective cohort of consecutive patients	rhBMP2/BGE n=41 (4.2-12 mg/pt)	single- or multi-level posterior cervical stenosis, ACDF nonunion, or unstable spondylosis	single- or multi-level instrumented posterior cervical spinal fusion with rhBMP2/BGE or ICBG	NR Iliac crest pain postharvest	NR	
		ICBG n=36					

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
Smucker et al., 2006 USA **Cervical Spine**	Retrospective case-control	rhBMP2/CRA n=69 (dose NR) CRA n=165	NR	single- or multi-level instrumented ACDF with rhBMP2/CRA or CRA alone	NR	NR	
Vaidya et al., 2007 USA **Cervical Spine**	Retrospective cohort of consecutive patients	rhBMP2 n=22 (1-3 mg/pt) ALG/DBM n=24	single- or multiple-level cervical DDD with radiculopathy or myelopathy	single- or multi-level primary instrumented ACDF with interbody fusion cages rhBMP2 on ACS or ALG/DBM	Oswestry Disability Index Mean score improvement (points) 0.5, 1.5, 3, 6, 12, 24 mos rhBMP2 -3.6, 6, 8, 8, 14, 24 Neck pain Mean score improvement (points) 0.5, 1.5, 3, 6, 12, 24 mos rhBMP2 2, 2, 2, 2, 3, 4 Arm pain Mean score improvement (points) 0.5, 1.5, 3, 6, 12, 24 mos rhBMP2 1, 1, 2, 2, 3, 4 Oswestry Disability Index Mean score improvement (points) 0.5, 1.5, 3, 6, 12, 24 mos ALG/DBM 2, 6, 10, 21, 28, 33 Neck pain Mean score improvement (points) 0.5, 1.5, 3, 6, 12, 24 mos ALG/DBM	NR	Both groups showed significant improvements from baseline, but there were no significant differences between groups in mean score or rates

Investigator (yr, country, ref #) Surgical Site	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Percent improved or success (p-value)	Comment
					4, 4, 4, 4, 5, 6 Arm pain Mean score improvement (points) 0.5, 1.5, 3, 6, 12, 24 mos ALG/DBM 3, 4, 3, 5, 5, 5		

Appendix Table 8. Comparative studies reported in the AHRQ HTA evaluating BMPs in spinal fusion: functional outcomes

Note. Abstraction tables copied directly from the AHRQ HTA report except that the references were changed to correspond to the current report. In addition, adverse events and complications were omitted as they were reported elsewhere.

Investigator (yr, country, ref #)	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score (p-value)	Outcome measure % improved or success (p-value)	Comment
On-label use							
Boden et al., 2000 USA **Lumbar spine**	Multicenter, nonblinded RCT	rhBMP2 (4.2-8.4 mg/pt) n=11	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	SF-36 physical function subscale Mean score improvement (points) 3, 6, 12. 24 mos rhBMP2 10, 18, 27, 38	Work status at 24 mos rhBMP2 10 of 11 (91%) pts working	No significant differences between groups
		ICBG n=3			ICBG 13, 27, 37, 37	ICBG 2 of 3 (67%)	
Burkus et al., 2002 USA **Lumbar spine**	Multicenter, nonblinded RCT	rhBMP2 (4.2-8.4 mg/pt) n=143	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	Median days return to work rhBMP2 64	Neurological status 1.5, 3, 6, 12, 24 mos rhBMP2 80, 84, 78, 82, 83	No significant differences between groups
						Work status 3, 6, 12, 24 mos rhBMP2 38, 51, 55, 66 working	
		ICBG n=136			ICBG 65	Neurological status 1.5, 3, 6, 12, 24 mos ICBG 84, 77, 81, 85, 84	
						Work status 3, 6, 12, 24 mos ICBG 28, 46, 50, 56 working	

Burkus et al., 2003 (Integrated analysis) **Lumbar spine** Note: may include pts in Burkus et al., 2003 ("Radio-graphic assessment …")	Retrospective combined comparative analysis	rhBMP2 n=277 (dose NR) ICBG n=402	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages	SF-36 physical component subscale Mean score improvement (points) pre, 3, 6, 12, 24 mos rhBMP2 9, 12, 14, 16 ICBG 5, 8, 10, 12 (p=0.0015, 0.0004, 0.0003, 0.0007)	Work status at 24 mos rhBMP2 103 (75%) who were working presurgery returned to work ICBG 109 (65%) who were working presurgery returned to work (p NSD)	rhBMP recipients returned to work a median 55 days sooner than ICBG graft recipients (adjusted p=0.0156)
Off-label use							
Boden et al., (2002) USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP2/CRM plus Texas Scottish Rite Hospital (TSRH) Spinal System (TSRHSS) n=11 (40 mg/pt) rhBMP2/CRM alone n=11 (40 mg/pt) ICBG plus TSRHSS n=5	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 ICBG	SF-36 physical component subscale Mean score improvement (points) 1.5, 3, 6, 17 mos rhBMP2/CRM/TSRHSS ~1, ~0, ~5, ~4 rhBMP2/CRM alone ~1, ~9, ~11, ~16 ICBG/TSRHSS ~1, ~3, ~2, ~17	NR	Both rhBMP2/CRM groups showed statistically significant improvements over baseline, the ICBG group did not
Burkus et al., (2005) USA **Lumbar Spine**	Multicenter, nonblinded RCT	rhBMP2 n=79 (8-12 mg/pt)	single-level lumbar lumbar DDD	primary single-level anterior lumbar fusion with a pair of threaded allograft	SF-36 physical component subscale Mean score improvement (points) 6, 12, 24 mos rhBMP2 14, 16, 15 Average days to return to work	NR	SF-36 scores in both groups showed steady improvement from 6 to 24 mos. postsurgery

Note: includes all pts from Burkus et al., 2002, rec# 11510; same pts as Burkus et al., 2006, rec# 6640		ICBG N=52		cortical bone dowels (CBD) plus rhBMP2 or ICBG	rhBMP2 89 SF-36 physical component subscale Mean score improvement (points) 6, 12, 24 mos ICBG 9, 11, 12 (p=0.001, 0.003, 0.015) Average days to return to work ICBG 96 (p=not significant)		**Spectrum corrected the SF-36 scores to reflect imrpvoement rather than raw scores**
Dawson et al., 2009 USA **Lumbar spine**	Multicenter nonblinded RCT	rhBMP2/CRM n=25 (12 mg/pt)	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	SF-36 physical component subscale Mean score improvement (points) 24 mos rhBMP2/CRM 13 SF-36 physical function subscale Mean score improvement (points) 24 mos rhBMP2/CRM 36	Work status at 24 mos rhBMP2/CRM 8 of 23 (3%5) working	The rhBMP2/CRM group appeared to improve faster than the ICBG group, but this impression was not statistically supported
		ICBG n=21			SF-36 physical component subscale Mean score improvement (points) 24 mos ICBG 10 SF-36 physical function subscale Mean score improvement (points) 24 mos ICBG 18	ICBG 6 of 20 (30%) working	
Dimar et al., (2009) USA **Lumbar**	Multicenter nonblinded RCT	rhBMP2/CRM n=239 (40 mg/pt)	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion	SF-36 physical component subscale Mean score improvement (points) 1.5, 3, 6, 12, 24 mos rhBMP2/CRM ~4, ~9, ~13, ~13, ~13	Work status at 24 mos rhBMP2/CRM 87 of 207 (42) working	SF-36 physical component scale mean score improvements at 24 mos. exceeded

Spine Note: contains pts in Glassman et al., 2007, rec# 4040; Dimar et al., 2006 rec# 5480; Glassman et al., 2005, rec# 8040		 ICBG n=224		plus rhBMP2 or ICBG	 ICBG ~4, ~8, ~9, ~10, ~10	 ICBG 89 of 184 (48) working	a 5.41 point threshold proposed to be clinically significant (Ware et al., 1994)
Glassman et al., (2008) USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP2/ACS n=50 (dose not reported) ICBG n=52	single- or multi-level lumbar DDD	single- or multi-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	SF-36 physical component subscale Mean score improvement (points) 3, 6, 12, 24 mos rhBMP2 7, 8, 10, 7 ICBG 7, 9, 10, 7	NR	Both groups showed substantial improvements over baseline, with no significant intergroup differences
Haid et al., (2004) USA **Lumbar Spine**	Multicenter, nonblinded RCT	rhBMP2 n=34 (4.2-8.4)	single-level lumbar DDD	single-level primary posterior lumbar interbody fusion (PLIF) interbody fusion cages plus rhBMP2 or ICBG	SF-36 physical component subscale Mean score improvement (points) 1.5, 3, 6, 12, 24 mos rhBMP2 ~5, ~10, ~12, ~14, ~14 Motor function Mean score improvement (points) 24 mos rhBMP2 4.5 Sensory function Mean score improvement (points) 24 mos rhBMP2 8.0 Reflex function Mean score improvement (points)	Overall neurological success 24 mos rhBMP2 100	Overall neurological success rate represents a combination of the four neurological measurements

					24 mos rhBMP2 7.0		
					Straight leg raise Mean score improvement (points) 24 mos rhBMP2 48		
					Median days to return to work rhBMP2 43		
		ICBG N=33			SF-36 physical component subscale Mean score improvement (points) 1.5, 3, 6, 12, 24 mos ICBG ~2, ~6, ~6, ~6, ~11	ICBG 100	
					Motor function Mean score improvement (points) 24 mos ICBG 2.8		
					Sensory function Mean score improvement (points) 24 mos ICBG 2.8		
					Reflex function Mean score improvement (points) 24 mos ICBG 5.4		
					Straight leg raise Mean score improvement (points) 24 mos ICBG 39		
					Median days to return to work		

					ICBG 137 (p=NSD)		
Glassman et al., (2007) USA **Lumbar Spine**	Retrospective with historical control group	rhBMP2 n=91 (12 mg/pt) ICBG n=35	single- and multi-level lumbar DDD, degenerative scoliosis, postdiscectomy instability, spinal stenosis, adjacent level degeneration	single- or multi-level primary or revision instrumented posterolateral lumbar fusion	NR	NR	Study only reported fusion data
Mummaneni et al., 2004 USA **Lumbar Spine**	Retrospective single-center cohort study	rhBMP2/AGB n=25 (8.4 mg/pt) ICBG N=19	single- or multi-level lumbar DDD	single- or multi-level primary transforaminal lumbar interbody fusion (TLIF) with interbody fusion cages with rhBMP2 plus AGB or ICBG alone	Prolo Scale Functional status subscale Mean score at F/U rhBMP2/AGB 3.8±0.9 ICBG 4.0±0.7	NR	No statistical analysis
Pradhan et al., 2006 USA **Lumbar Spine**	Prospective consecutive patient single-center cohort study	rhBMP2 n=9 (dose NR) ICBG n=27	single-level lumbar DDD	single-level primary anterior lumbar interbody fusion (ALIF) with femoral ring allograft	NR	NR	Study only reported fusion data

				(FRA) plus rhBMP2 or ICBG			
Singh et al., 2006 USA **Lumbar Spine**	Prospective single-center case-matched cohort study	rhBMP2/ICBG n=39 (12-36 mg/pt) ICBG N=11	single- or multi-level lumbar DDD	single- or multi-level primary instrumented posterolateral lumbar fusion with rhBMP2 plus ICBG or ICBG alone	NR	NR	
Slosar et al., 2007 USA **Lumbar Spine**	Prospective consecutive patient single-center cohort study	rhBMP2 n=45 (3-9 mg/pt) ALG N=30	single- or multi-level lumbar DDD	single- or multi-level primary instrumented anterior lumbar interbody fusion (ALIF) with femoral ring allograft (FRA) plus rhBMP2 or allograft bone chips (ALG)	NR	NR	
Johnsson et al., 2002 Sweden **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP7 n=10 (7 mg/pt) ICBG n=10	single-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	NR	NR	
Kanayama et al., 2006 Japan, USA	Multicenter nonblinded RCT	rhBMP7 n=9 (7 mg/pt)	single-level lumbar DDD	single-level primary instrumented posterolateral	NR	NR	

Lumbar Spine		AGB/CRM n=10		lumbar fusion with rhBMP7 or AGB/CRM			
Vaccaro et al., 2008 USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP7 n=207 (7 mg/pt) ICBG n=86	single-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	NR	Neurological success 36+ mos rhBMP7 84 ICBG 80	Neurological success is a composite outcome comprising muscle strength, reflexes, sensation, and straight leg raise
Vaccaro et al., 2008 USA **Lumbar Spine** Note: Long-term F/U study that includes all pts from Vaccaro et al., 2004, and Vaccaro et al., 2005	Multicenter, nonblinded RCT	rhBMP7 n=24 (7 mg/pt) ICBG n=12	single-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	NR	Patients in both groups displayed increases in the SF-36 physical component subscale, increasing from the 25th percentile, reaching age-matched normative values at 48 mos. (data not shown)	
Baskin et al., 2003 USA **Cervical Spine**	Multicenter, nonblinded RCT	rhBMP2/ALG n=18 (0.6-1.2 mg/pt)	single- or two-level cervical DDD	single- or two-level primary instrumented ACDF with rhBMP2/ALG or ICBG/ALG	SF-36 physical component subscale Mean score improvement (points) 1.5, 3, 6, 12, 24 mos rhBMP2/ALG 9, 13, 14, 14, 17 SF-36 mental component subscale Mean score improvement (points) 1.5, 3, 6, 12, 24 mos rhBMP2/ALG 19, 16, 22, 22, 22	SF-36 physical component subscale 24 mos rhBMP2/ALG 92 SF-36 mental component subscale 24 mos rhBMP2/ALG 92	No significant differences between group

						Neurological status 1.5, 3, 6, 12, 24 mos rhBMP2/ALG 94, 100, 88, 100, 100	
		ICBG/ALG n=15			SF-36 physical component subscale Mean score improvement (points) 1.5, 3, 6, 12, 24 mos ICBG/ALG 7, 12, 14, 16, 16	SF-36 physical component subscale 24 mos ICBG/ALG 100	
					SF-36 mental component subscale Mean score improvement (points) 1.5, 3, 6, 12, 24 mos ICBG/ALG 10, 5, 12, 8, 7	SF-36 mental component subscale 24 mos ICBG/ALG 75	
						Neurological status 1.5, 3, 6, 12, 24 mos ICBG/ALG 100, 100, 100, 93, 100	
Butterman et al., 2008 USA **Cervical Spine**	Prospective nonrandomized cohorts of consecutive patients	rhBMP2/CRA n=30 (0.9-3.7 mg/pt)	single- or multiple-level cervical DDD	single- or multi-level primary instrumented or uninstrumented ACDF with rhBMP2/CRA or ICBG	NR	Resolution of neurological deficits manifested as weakness and altered sensation rhBMP2/CRA 100	
		ICBG n=36				ICBG 100	
Crawford et al., 2009 USA **Cervical Spine**	Retrospective cohort of consecutive patients	rhBMP2/BGE n=41 (4.2-12 mg/pt)	single- or multi-level posterior cervical stenosis, ACDF nonunion, or unstable spondylosis	single- or multi-level instrumented posterior cervical spinal fusion with rhBMP2/BGE or ICBG	NR	NR	
		ICBG n=36					
Smucker et	Retrospective	rhBMP2/CRA	NR	single- or	NR	NR	

al., 2006 USA **Cervical Spine**	case-control	n=69 (dose NR) CRA n=165		multi-level instrumented ACDF with rhBMP2/CRA or CRA alone			
Vaidya et al., 2007 USA **Cervical Spine**	Retrospective cohort of consecutive patients	rhBMP2 n=22 (1-3 mg/pt) ALG/DBM n=24	single- or multiple-level cervical DDD with radiculopathy or myelopathy	single- or multi-level primary instrumented ACDF with interbody fusion cages rhBMP2 on ACS or ALG/DBM	NR	NR	

Appendix Table 9. Comparative studies reported in the AHRQ HTA evaluating BMPs in spinal fusion: quality of life and patient satisfaction outcomes

Note. Abstraction tables copied directly from the AHRQ HTA report except that the references were changed to correspond to the current report. In addition, adverse events and complications were omitted as they were reported elsewhere.

Investigator (yr, country, ref #)	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score	Outcome measure % improved or success (p-value)	Comment
On-label use							
Boden et al., 2000 USA **Lumbar spine**	Multicenter, nonblinded RCT	rhBMP2 (4.2-8.4 mg/pt) n=11 ICBG n=3	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	SF-36 general health perception subscale Mean score 0, 3, 6, 12, 24 mos rhBMP2 68, 74, 68, 70, 73 ICBG 59, 57, 75, 64, 67	All improved over 24 mos. (p not reported)	At 24 mos. 11 of 11 pts in rhBMP2 group rated outcome as excellent; 1 of controls rated outcome as excellent, 1 each good and fair. Mean neurologic scores were increased over baseline at all time points in both groups.
Burkus et al., 2002 USA **Lumbar spine**	Multicenter, nonblinded RCT	rhBMP2 (4.2-8.4 mg/pt) n=143 ICBG n=136	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages plus rhBMP2 or ICBG	NR	Patient satisfaction 24 mos rhBMP2 81% satisfied ICBG 80% satisfied	82% of rhBMP group indicated they would undergo same procedure, compared with 77% of ICBG group
Burkus et al., 2003 (Integrated analysis) **Lumbar spine** Note: may include pts in Burkus et al., 2003 (“Radio-	Retrospective combined comparative analysis	rhBMP2 n=277 (dose NR) ICBG n=402	single-level lumbar DDD	single-level primary anterior lumbar fusion with interbody fusion cages	NR	NR	

Investigator (yr, country, ref #)	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score	Outcome measure % improved or success (p-value)	Comment
graphic assessment…")							
Off-label use							
Boden et al., (2002) USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP2/CRM plus Texas Scottish Rite Hospital (TSRH) Spinal System (TSRHSS) n=11	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 ICBG	NR	Patient satisfaction (% good/excellent) pre, 1.5, 3, 6, 17 mos rhBMP2/CRM/TSRHSS 0, ~75, ~58, ~60, ~60	Patient satisfaction measurements generally paralleled results of SF-36 pain survey and Oswestry DI
						Physician impression (% good/excellent) pre, 1.5, 3, 6, 17 mos rhBMP2/CRM/TSRHSS 0, ~90, ~80, ~80, ~80	
		(40 mg/pt) rhBMP2/CRM alone n=11				Patient satisfaction (% good/excellent) pre, 1.5, 3, 6, 17 mos rhBMP2/CRM alone 0, ~100, ~88, ~88, ~100	
						Physician impression (% good/excellent) pre, 1.5, 3, 6, 17 mos rhBMP2/CRMalone 0, ~100, ~85, ~80, ~85	
		(40 mg/pt) ICBG plus TSRHSS n=5				Patient satisfaction (% good/excellent) pre, 1.5, 3, 6, 17 mos ICBG/TSRHSS 0, ~80, ~60, ~80, ~60	
						Physician impression (% good/excellent) pre, 1.5, 3, 6, 17 mos ICBG/TSRHSS	

Investigator (yr, country, ref #)	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score	Outcome measure % improved or success (p-value)	Comment
						0, ~60, ~80, ~60, ~60	
Burkus et al., (2005) USA **Lumbar Spine** Note: includes all pts from Burkus et al., 2002, rec# 11510; same pts as Burkus et al., 2006, rec# 6640	Multicenter, nonblinded RCT	rhBMP2 n=79 (8-12 mg/pt) ICBG N=52	single-level lumbar lumbar DDD	primary single-level anterior lumbar fusion with a pair of threaded allograft cortical bone dowels (CBD) plus rhBMP2 or ICBG	NR	NR	
Dawson et al., 2009 USA **Lumbar spine**	Multicenter nonblinded RCT	rhBMP2/CRM n=25 (12 mg/pt) ICBG n=21	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	NR	NR	
Dimar et al., (2009) USA **Lumbar Spine** Note: contains pts in Glassman et al., 2007, rec# 4040; Dimar et al., 2006 rec# 5480; Glassman et al., 2005, rec# 8040	Multicenter nonblinded RCT	rhBMP2/CRM n=239 (40 mg/pt) ICBG n=224	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG	NR	NR	
Glassman et al., (2008)	Multicenter nonblinded	rhBMP2/ACS n=50	single- or multi-level	single- or multi-level primary	NR	NR	

Investigator (yr, country, ref #)	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score	Outcome measure % improved or success (p-value)	Comment
USA **Lumbar Spine**	RCT	(dose not reported) ICBG n=52	lumbar DDD	instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG			
Haid et al., (2004) USA **Lumbar Spine**	Multicenter, nonblinded RCT	rhBMP2 n=34 (4.2-8.4) ICBG N=33	single-level lumbar DDD	single-level primary posterior lumbar interbody fusion (PLIF) interbody fusion cages plus rhBMP2 or ICBG		Patient satisfaction at 24 mos rhBMP2 72 ICBG 80	Patient satisfaction rates comprise results for pts who report definitely and mostly true that they were satisfied with their surgical outcomes
Glassman et al., (2007) USA **Lumbar Spine**	Retrospective with historical control group	rhBMP2 n=91 (12 mg/pt) ICBG n=35	single- and multi-level lumbar DDD, degenerative scoliosis, postdiscectomy instability, spinal stenosis, adjacent level degeneration	single- or multi-level primary or revision instrumented posterolateral lumbar fusion	NR	NR	Study only reported fusion data
Mumma-neni et al., 2004 USA **Lumbar Spine**	Retrospective single-center cohort study	rhBMP2/AGB n=25 (8.4 mg/pt)	single- or multi-level lumbar DDD	single- or multi-level primary transforaminal lumbar interbody fusion (TLIF) with interbody fusion cages with	Prolo Scale Economic status subscale Mean score at F/U rhBMP2/AGB 3.8±0.8 Medication use subscale	NR	Statistical analysis not done

Investigator (yr, country, ref #)	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score	Outcome measure % improved or success (p-value)	Comment
		ICBG N=19		rhBMP2 plus AGB or ICBG alone	Mean score at F/U rhBMP2/AGB 3.8±0.9 Prolo Scale Economic status subscale Mean score at F/U ICBG 4.1±0.7 Medication use subscale Mean score at F/U ICBG 4.2±0.8		
Pradhan et al., 2006 USA **Lumbar Spine**	Prospective consecutive patient single-center cohort study	rhBMP2 n=9 (dose NR) ICBG n=27	single-level lumbar DDD	single-level primary anterior lumbar interbody fusion (ALIF) with femoral ring allograft (FRA) plus rhBMP2 or ICBG	NR	NR	Study only reported fusion data
Singh et al., 2006 USA **Lumbar Spine**	Prospective single-center case-matched cohort study	rhBMP2/ICBG n=39 (12-36 mg/pt) ICBG N=11	single- or multi-level lumbar DDD	single- or multi-level primary instrumented posterolateral lumbar fusion with rhBMP2 plus ICBG or ICBG alone	NR	NR	
Slosar et al., 2007	Prospective	rhBMP2	single- or	single- or multi-	NR	Patient satisfaction at 24	None of the pts who

Investigator (yr, country, ref #)	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score	Outcome measure % improved or success (p-value)	Comment
USA **Lumbar Spine**	consecutive patient single-center cohort study	n=45 (3-9 mg/pt) ALG N=30	multi-level lumbar DDD	level primary instrumented anterior lumbar interbody fusion (ALIF) with femoral ring allograft (FRA) plus rhBMP2 or allograft bone chips (ALG)		mos rhBMP2 86 ALG 79	underwent revision fusions in ALG group expressed satisfaction with their outcomes
Johnsson et al., 2002 Sweden **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP7 n=10 (7 mg/pt) ICBG n=10	single-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	NR	NR	
Kanayama et al., 2006 Japan, USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP7 n=9 (7 mg/pt) AGB/CRM n=10	single-level lumbar DDD	single-level primary instrumented posterolateral lumbar fusion with rhBMP7 or AGB/CRM	NR	NR	
Vaccaro et al., 2008 USA **Lumbar Spine**	Multicenter nonblinded RCT	rhBMP7 n=207 (7 mg/pt) ICBG n=86	single-level lumbar DDD	single-level primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG	NR	NR	
Vaccaro et al., 2008	Multicenter,	rhBMP7	single-level	single-level	NR	Patients in both groups	

Investigator (yr, country, ref #)	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score	Outcome measure % improved or success (p-value)	Comment
USA **Lumbar Spine** Note: Long-term F/U study that includes all pts from Vaccaro et al., 2004, and Vaccaro et al., 2005	nonblinded RCT	n=24 (7 mg/pt) ICBG n=12	lumbar DDD	primary uninstrumented posterolateral lumbar fusion with rhBMP7 or ICBG		displayed increases in the SF-36 mental health component subscale, increasing from the 25th percentile, reaching age-matched normative values at 48 mos. (data not shown)	
Baskin et al., 2003 USA **Cervical Spine**	Multicenter, nonblinded RCT	rhBMP2/ALG n=18 (0.6-1.2 mg/pt) ICBG/ALG n=15	single- or two-level cervical DDD	single- or two-level primary instrumented ACDF with rhBMP2/ALG or ICBG/ALG	NR	Patient satisfaction 24 mos > 90% in both groups	Patient satisfaction related to whether they were satisfied with their results, whether they were helped as much as anticipated, and whether they would have the surgery again
Butterman et al., 2008 USA **Cervical Spine**	Prospective nonrandomized cohorts of consecutive patients	rhBMP2/CRA n=30 (0.9-3.7 mg/pt) ICBG n=36	single- or multiple-level cervical DDD	single- or multi-level primary instrumented or uninstrumented ACDF with rhBMP2/CRA or ICBG	NR	Patient-reported success 13-24, 25-36 mos rhBMP2/CRA 90, 89 ICBG 94, 97	Patient satisfaction related to whether they were satisfied with their results, whether they would have the surgery again, and whether they would recommmend ot to others (97% in both groups)
Crawford et al., 2009 USA	Retrospective cohort of consecutive patients	rhBMP2/BGE n=41 (4.2-12 mg/pt) ICBG	single- or multi-level posterior cervical	single- or multi-level instrumented posterior	NR	NR	

Investigator (yr, country, ref #)	Study design	Comparisons No. pts (BMP dose)	Patient diagnosis	Surgical intervention	Outcome measure mean score	Outcome measure % improved or success (p-value)	Comment
Cervical Spine		n=36	stenosis, ACDF nonunion, or unstable spondylosis	cervical spinal fusion with rhBMP2/BGE or ICBG			
Smucker et al., 2006 USA **Cervical Spine**	Retrospective case-control	rhBMP2/CRA n=69 (dose NR) CRA n=165	NR	single- or multi-level instrumented ACDF with rhBMP2/CRA or CRA alone	NR	NR	
Vaidya et al., 2007 USA **Cervical Spine**	Retrospective cohort of consecutive patients	rhBMP2 n=22 (1-3 mg/pt) ALG/DBM n=24	single- or multiple-level cervical DDD with radiculopathy or myelopathy	single- or multi-level primary instrumented ACDF with interbody fusion cages rhBMP2 on ACS or ALG/DBM	NR	NR	

Appendix Table 10. Comparative studies reported in the AHRQ HTA evaluating BMPs in spinal fusion: detailed results

Investigator	Outcomes mean ± SD (unless otherwise indicated) (range)			
Off-label use				
Carragee et al	**NONE (only safety – retrograde ejaculation**			
Crawford et al (2010) Retrospective cohort **Sacrum** Treatment groups: *rhBMP*: n = 36 *autograft*: n = 24 (historical controls) (appears to contain the same patients reported in Maeda (2009) [60])	**Surgical outcomes**			
	Surgical time (h ± SD): *rhBMP2*: 10.8 ± 2.5 *Autograft*: 11.3 ± 3.0 *P* = ns	Estimated blood loss (mL ± SD): *rhBMP2*: 1221 ± 903 *Autograft*: 1938 ± 1190 *P* = .007	Spinal osteotomy: *rhBMP2*: 39.8% (14/36) *Autograft*: 50.0% (12/24) Posterior fusion only: *rhBMP2*: 11.1% (4/36) *Autograft*: 0% (0/24)	New levels fused: *rhBMP2*: 2.6 ± 1.7 *Autograft*: 2.6 ± 1.8 *P* = ns
	Radiographic outcomes			
	Successful outcome (fusion grade 1 or 2) rhBMP: 88.9% (32/36) Autograft: 79.2% 19/24)	Fusions evaluated using a 4-point scale: grade 1: definite fusion; grade 2: probable fusion; grade 3: probable nonunion; grade 4: definite nonunion (pseudarthrosis) Where differences existed between the 2 evaluators, the average for the region was calculated and used for final analysis. Pseudarthrosis was defined as a fusion mass with a grade 3 or 4 or by the presence of implant failure (broken rods, broken screws, disengaged rods, screw loosening at bone implant interface) consistent with previously published pseudarthrosis criteria.	Posterior fusion grade from L4 to the sacrum *rhBMP*: 1.7 ± 0.9 *Autograft*: 2.3 ± 0.7 *P* = .021 (significantly better fusion in the rhBMP group)	
	Pain			
	ODI *rhBMP2*: preop: 38.5 ± 11.7 final postop: 20.1 ± 13.1 improvement: 18.4 *autograft*:	"Success" not reported/defined	Scoliosis Research Society (SRS-30) Pain *rhBMP2*: preop: 2.8 ± 0.6 final postop: 3.8 ± 0.7 improvement: 1.0 ± 0.7	"Success" not reported/defined

	preop: 44.8 ± 22.2 final postop: 22.5 ± 19.5 improvement: 22.3 *P* = ns		*Autograft*: preop: 2.7 ± 1.3 final postop: 3.9 ± 0.9 Improvement: 1.2 ± 0.9 *P* = ns	
	Function			
	SRS Activity domain *rhBMP2*: preop: 3.1 ± 0.5 final postop: 3.7 ± 0.5 improvement: 0.6 ± 0.5 *autograft*: preop: 2.8 ± 0.9 final postop: 3.7 ± 0.6 improvement: 0.9 ± 0.8 *P* = ns	"Success" not reported/defined		
	Work status			
	NR			
	Neurological status			
	NR			
	Social function & mental health			
	SRS Self-image domain *rhBMP2*: preop: 2.7 ± 0.7 final postop: 3.7 ± 0.8 improvement: 1.0 ± 0.9 *autograft*: preop: 2.6 ± 0.9 final postop: 3.4 ± 0.7 improvement: 0.8 ± 0.7 *P* = ns	*SRS mental health domain* *rhBMP2*: preop: 3.7 ± 0.7 final postop: 4.0 ± 0.7 improvement: 0.3 ± 0.7 *autograft*: preop: 2.3 ± 1.8 final postop: 3.8 ± 0.8 improvement: 1.5 *P* = NR		
	Patient Satisfaction			
	SRS Satisfaction (final score) At Final follow-up *rhBMP2*: 4.2 ± 0.9 *Autograft*: 4.0 ± 0.7 *P* = ns			
Howard et al. (2011)	**NONE (only safety – graft site pain)**			
Latzman et al. (2010) rhBMP2: n = 24 Auto/allograft n = 105	**Surgical outcomes**			
	Packed RBC transfusion intraoperatively rhBMP2: 25.9% (7/27) Auto/allograft: 9.3% (10/108)			
	Radiographic, Pain, Function			
	NR			
Lee et al. (2010)	**Surgical outcomes**			
	NR			

Retrospective cohort **Lumbar spine** rhBMP2 age ≥ 65 years: n = 34 rhBMP2 age < 65 years: n = 52 ICBG age ≥ 65 years: n = 41	**Radiographic outcomes**			
	rhBMP2 age ≥ 65 years vs. < 65 years: **Fusion rate** 82.4% (28/34) vs. 94.2% (49/52) *P* = ns **Noticed fusion time** 95.7 ± 24.4 days vs. 83.7 ± 32.5 days *P* = .01 **Solid fusion time** 259.1 ± 76.9 days vs. 248.3 ± 77.3 days *P* = ns ***rhBMP2 vs. ICBG (age ≥ 65 years):*** **Fusion rate** 82.4% (28/34) vs. 78.1% (32/41) *P* = ns **Noticed fusion time** 95.7 ± 24.4 days vs. 102.5 ± 24.5 days *P* = ns **Solid fusion time** 259.1 ± 76.9 days vs. 291.8 ± 68.8 days *P* = ns ***Multivariable analysis of patients age ≥ 65 years with rhBMP2 vs. ICBG:*** **Fusion rate** *Females*: 87.5% vs. 79.2%; *Multilevel fusion*: 82.4% vs. 75.0%; *Smokers*: 60.0% vs. 57.1%; *Osteoporosis*: 85.7% vs. 77.8%; *Post-revision*: 83.4% vs. 100%; *Multiple comorbidities*: 77.8% vs. 83.4% *P* = ns for all comparisons **Noticed fusion time** *Females*: 98.1 ± 21.3 vs. 105.5 ± 26.6 days; *Multilevel fusion*: 100.4 ± 22.9 vs. 97.5 ± 17.2 days; Smokers: 121.1 ± 32.3 vs. 127.6 ± 33.5 days; *Osteoporosis*: 98.5 ± 17.1 vs. 103.5 ± 21.1 days; Post-revision: 95.1 ± 27.6 vs. 101.8 ± 24.2 days; *Multiple comorbidities*: 103.6 ± 19.8 vs. 103.5 ±	"Success" not reported/defined	*Noticed fusion* = the first presence of bridging bone between two transverse processes in the fusion segment; *Solid fusion* = the clear presence of a robust fusion mass with consolidated bridging bone.	

	25.1 days *P* = ns for all comparisons **Solid fusion time** *Females*: 256.8 ± 71.8 vs. 285.5 ± 66.7 days; *Multilevel fusion*: 293.2 ± 61.9 vs. 294.1 ± 62.6 days; *Smokers*: 295.7 ± 99.6 vs. 319.6 ± 76.9 days; *Osteoporosis*: 279.5 ± 72.2 vs. 287.4 ± 59.7 days; *Post-revision*: 256.8 ± 71.8 vs. 256.8 ± 71.8 days; *Multiple comorbidities*: 299.9 ± 70.6 vs. 289.9 ± 69.4 days *P* = ns for all comparisons			
	Pain			
	VAS pain scores (0-10) *rhBMP2 age ≥ 65 years:* preop: 7.8 6 months: 2.8 1 year: 3.4 2 years: 4.1 rhBMP2 age < 65 years: preop: 7.7 6 months: 3.0 1 year: 3.1 2 years: 3.3 ICBG age ≥ 65 years: preop: 7.8 6 months: 2.9 1 year: 3.3 2 years: 3.9 *P* = .04 at 2 years between rhBMP2 age ≥ 65 years and age < 65 years	"Success" not reported/defined		
	Function			
	rhBMP2 age ≥ 65 years vs. age < 65 years: *'Good' outcome (Kirkaldy-Willis):* 85.3% (29/34) vs. 92.3% (48/52) *P* = ns	**rhBMP2 age ≥ 65 years vs. ICBG age ≥ 65 years:** *'Good' outcome (Kirkaldy-Willis):* 85.3% (29/34) vs. 73.2% (30/41) *P* = ns	Clinical outcomes were assessed based on a 4-grade system (Kirkaldy-Willis): 'excellent', 'good', 'fair' and 'poor'. 'Good' and 'excellent' were further classified as good results, and 'fair' and 'poor' were further classified as poor results	
Taghavi et al. (2010) Retrospective cohort **Lumbar spine** rhBMP2 n = 24	**Surgical outcomes**			
	NR			
	Radiographic outcomes			
	Fusion rate *Overall* rhBMP2: 100% (24/24) BMAA: 77.8% (14/18) Autograft: 100% (20/20) *P* = .01 for rhBMP2 and Autograft vs. BMAA	**Time to Solid Fusion (days)** *Overall* rhBMP2: 218.4 ± 63.8 BMAA: 297.6 ± 68.3 Autograft: 270.0 ± 60.4	3 criteria were used for assessment of fusion: (1) the presence of trabeculated bone between transverse processes, (2) no implant loosening and	

BMAA N = 18 Autograft N = 20	*Single-level* rhBMP2: 100% (13/13) BMAA: 100% (7/7) Autograft: 100% (10/10) *P* = ns *Mutilevel* rhBMP2: 100% (11/11) BMAA : 63.6% (7/11) Autograft: 100% (10/10) *P* = .02 for rhBMP2 and Autograft vs. BMAA	*P* = .002 and .03 for rhBMP2 group vs. BMAA and Autograft, respectively *Single-level* rhBMP2: 199.8 ± 49.8 BMAA: 313.3 ± 34.3 Autograft: 276.7 ± 29.8 *P* = .001 and < .001 for rhBMP2 group vs. Autograft and BMAA, respectively *Mutilevel* rhBMP2: 240.4 ± 71.3 BMAA : 282.0 ± 87.5 Autograft: 263.3 ± 79.4 *P* = ns for all comparisons	(3) less than 2° of movement on lateral flexion and extension films. A diagnosis of nonunion was based on exploration during an additional revision surgery or evidence of nonunion on dynamic radiographs or computerized tomography.	
	Pain			
	VAS back pain (0-10)* *rhBMP* preop: 8.2 1.5 mos.: 3.3 6 mos.: 3.7 1 year: 3.6 2 years: 3.9 *BMAA* preop: 8.2 1.5 mos.: 4.0 6 mos.: 4.2 1 year: 4.2 2 years: 4.3 *Autograft* preop: 7.9 1.5 mos.: 3.5 6 mos.: 3.6 1 year: 3.9 2 years: 3.9 *P* < .001 for decrease in preop and 2-year scores in all groups; no significant differences seen between groups at any time point.	**VAS leg pain (0-10)*** *rhBMP* preop: 7.9 1.5 mos.: 2.9 6 mos.: 3.4 1 year: 3.4 2 years: 3.6 *BMAA* preop: 7.9 1.5 mos.: 3.6 6 mos.: 3.9 1 year: 3.8 2 years: 3.9 *Autograft* preop: 7.7 1.5 mos.: 3.0 6 mos.: 3.4 1 year: 3.5 2 years: 3.6 *P* < .001 for decrease in preop and 2-year scores in all groups; no significant differences seen between groups at any time point.		
	Function			
	NR			
Delawi et al. (2010) RCT **Lumbar spine** Treatment groups: *rhOP-1*: n = 18	**Surgical outcomes**			
	Surgical time (min ± SD): *rhOP-1*: 178 ± 73 *Autograft*: 178 ± 47 *P* = ns	Estimated blood loss (mL ± SD): *rhOP-1*: 422 ± 265 *Autograft:* 373 ± 301 *P* = ns	Hospital stay (day ± SD) *rhOP-1*: 10.5 ± 4.9 *Autograft:* 10.9 ± 6.4 *P* = ns	
	Radiographic outcomes			

autograft: n = 16	**Fusion rates on CT at 1 year:** Definite fusion: *OP-1*: 62.5% (10/16) *Autograft*: 66.7% (10/15) Doubtful fusion: *OP-1*: 25.0% (4/16) *Autograft*: 20.0% (3/15) Nonunion: *OP-1*: 12.5% (2/16) *Autograft*: 13.3% (2/15) *P* = ns for all comparisons	Successful outcome (definite fusion) *OP-1*: 62.5% (10/16) *Autograft*: 66.7% (10/15) *P* = ns	Fusion classified via system of *Christensen et al* (3 categories): 1. "Fusion" = a continuous bony bridge from the base of the pedicle and transverse processes from 1 vertebra to the other, at a minimum of 1 side of the spine, in absence of any secondary signs of nonunion, such as fracture or loosening of the screws. If the fusion was doubtful in any way, the patient was not classified as fused. 2. "Doubtful fusion" = suboptimal quality of the bone bridging or some doubtful discontinuity, including fusion mass possibly hidden behind instrumentation, at a minimum of 1 side of the spine, in the absence of "fusion" on the other side. 3. "Nonunion" = definite discontinuity or lack of the fusion mass at both sides of the spine.	CT scans were reviewed by a spinal surgeon and a senior radiology resident blinded to the treatment group and the institute where the procedure was performed. A third observer, a spinal surgeon, was used to adjudicate conflicting findings. In the exceptional case that all 3 observers classified the fusion differently, the patient was classified as "Doubtful fusion."
	Function/ADLs			
	Mean ODI scores (OP-1 vs. Autograft) *Preop:* 44 ± 15 vs. 53 ± 13 *6 weeks†:* 33 vs. 47 *3 months†:* 17 vs. 35 *6 months†:* 20 vs. 30 *12 months†:* 17 vs. 29 *P* = ns for between group comparisons all time points; P < .001 for scores at all follow-up time points compared with preop for both groups.	% Success NR		
Hwang et al. (2010) RCT	**NONE (only safety & special populations)**			
Cahill et al. (2011) Retrospective	**Surgical outcomes**			
	Surgical time (min ± SD): NR	Estimated blood loss (mL):	Hospital stay (day ± SD)	

case-control (database) study **Lumbar spine** Treatment groups: *rhBMP (any)*: n = 2372 *Non-BMP*: n = 2372		*NR*	*BMP:* 3 days (median) *No BMP:* 3 days (median) (*P* = .5)	
Xu et al. (2011) Retrospective cohort study **Cervical spine** Treatment groups: *rhBMP-2*: n = 48 *Non-BMP*: n = 156	**Surgical outcomes**			
	Surgical time (min ± SD): NR	Estimated blood loss (mL): *rhBMP2*: 500 (range, 200, 700) *Non-BMP:* 300 (range, 200, 425) *P* = .45	Hospital stay (day ± SD) *rhBMP2*: 6.1 ± 4.7 *Non-BMP:* 7.4 ± 6.9 *P* = .23	
	Radiographic outcomes			
	Fusion rates on plain radiographs and CT at last f/u (>6 mos only) Fusion: *rhBMP2*: 100% (48/48) *Non-BMP:* 87.6% (106/121) *P* = .01			
	Pain			
	Neck pain (at last f/u): *rhBMP2*: 48% (19/48) *Non-BMP:* 23.3% (31/156) *P* = .003	% Success NR	NR	
	Function/ADLs			
	Nurick score (mean ± SD) Baseline: *rhBMP2*: 2.37 ± 1.51 *Non-BMP:* 2.51 ± 1.36 *P* = .11 Last f/u (24.2 ± 10.1 mos): *rhBMP2*: 1.30 ± 1.15 *Non-BMP:* 1.34 ± 1.49 *P* = .61	**ASIA score (mean ± SD)** Baseline: *rhBMP2*: 4.02 ± 0.68 *Non-BMP:* 3.88 ± 0.75 *P* = .10 Last f/u (24.2 ± 10.1 mos): *rhBMP2*: 4.39 ± 0.80 *Non-BMP:* 4.39 ± 0.78 *P* = .96		
Yaremchuck et al. (2010)[58] Retrospective case-control (database) study **Lumbar spine**	**Surgical outcomes**			
	Surgical time (min ± SD): NR	Estimated blood loss (mL): *NR*	Hospital stay (day ± SD) Total LOS: *BMP:* 8.4 ± 7.25 days *No BMP:* 5.5 ± 4.5 days (*P* = NR)	

Treatment groups: *BMP (any)*: n = 2372 *Non-BMP*: n = 2372			LOS before surgery: *BMP:* 1.2 ± 3.4 days *No BMP:* 1.2 ± 3.8 days (*P* = .859) LOS after surgery: *BMP:* 7.2 ± 11.1 days *No BMP:* 4.3 ± 5.2 days (*P* = .001)	

ADLs: activities of daily living; IQR: interquartile range; LBP: low back pain; NR: not reported; ODI: Oswestry Low Back Pain Disability Questionnaire; OR: odds ratio; SLR: straight leg raise; SRS: Scoliosis Research Society
*Means estimated from graphs/figures provided in the article.
†Based on the number of operations: rhBMP2, n = 27; auto/allograft, n = 108.
‡Adjusted for age, race, sex, income, elective admission, teaching hospital, revision surgery, diagnosis, medical comorbidities, levels fused, primary payer, and geographic location of hospital.

Appendix Table 11. Safety data from comparative studies

Investigator	Surgical and perioperative complications	Adverse events	Second surgeries	Iliac crest graft site
On-label use				
Boden (2000) (AHRQ ref 71) RCT pilot study **Lumbar spine** **On-label** Single-level primary anterior fusion with interbody fusion cages **rhBMP2 (n = 11) vs. ICBG (n = 3)**	*rhBMP2 vs. ICBG* ***Bowel obstruction*** (postop) & delay in gait training: 9% (1/11) vs. 33% (1/3) ***Wound dehiscence:*** 9% (1/11) vs. 0% (0/3)	*rhBMP2 vs. ICBG* ***Urinary retention:*** 0% (0/11) vs. 33% (1/3) ***Graft subsidence*** 0% (0/11) vs. 0% (0/3) ***Graft migration:*** 0% (0/11) vs. 0% (0/3) ***Graft rotation:*** 0% (0/11) vs. 0% (0/3) ***Episode of LBP:*** 9% (1/11) (prior to 6 mos. f/u) vs. 0% (0/3) ***Postoperative traumatic events:*** 27% (3/11) (falls) vs. 0% (0/3) ***Deaths (cumulative)*** *(not attributed to treatment):* 0% (0/123) vs. 0.1% (1/109) (cause NR; death occurred between 6-12 mos. f/u) Blood tests showed no differences in CBC or blood chemistry Elevated rhBMP2 antibody titers: 0% (0/11) Antibovine collagen antibodies: 27% (3/11) (no clinical sequelae).	*rhBMP2:* 0% (0/11) *ICBG:* 33% (1/3)-pseudoarthroses; supplemental posterolateral instrumented fusion at 18 mos.	NR
Burkus (2002) (AHRQ ref 72) RCT **Lumbar spine** **On-label** Single-level primary anterior fusion with interbody fusion cages **rhBMP2 (n = 143) vs.**	<u>Surgical & perioperative complications</u> "There were no unanticipated [surgical] device-related adverse events in either treatment group." ***Vascular events:*** 4.2% (6/143) vs. 3.7%	<u>Adverse events</u> *rhBMP2 vs. ICBG* ***Retrograde ejaculation (RE):*** 4.1% (6/146) of all males (tx group NR) (postsurgical) -permanent RE: (4/146) 2.8% (tx group NR) ***Implant displacement:***	<u>Second surgeries</u> *rhBMP2:* 7.7% (11/143) -implant removals (2/143) (5 days due to vertebral bone fracture and implant displacement; 4 mos. due to implant displacement and	<u>Iliac crest graft site</u> Any adverse event: 5.9% (8/136): -injury to lateral femoral cutaneous nerve: 2.2% (3/136) -avulsion fractures of anterior superior iliac crest: 1.5% (2/136) -infection (superficial): (0.7% (1/136)

ICBG (n = 136)	(5/136) -Laceration of iliac vein 6/279 (tx group NR) -deep vein thrombosis: 0% (0/143) vs. 1.5% (2/136)	see second surgeries ***Pseudoarthrosis:*** see second surgeries Elevated rhBMP2 antibody titers: 0.7% vs. 0.8% (3 mos.) (no negative consequences) ***Antibovine collagen antibodies***: 27% (3/11) (2 transient, 1 persistent but the patient had a positive titer prior to surgery. No correlation with clinical outcomes). <u>(as reported in radiographic results):</u> ***Radiolucencies*** from micromotion at implant-host bone interface increased over time (both groups, data NR). ***Atrophy of bone grafts*** over time (data NR).	possible failed fusion) - supplemental fixation for pseudoarthrosis (7/143) (all 7 seven patients had radiographically solid fusion but repeat surgery done due to persistent pain) -supplemental fixation after posterior decompression for persistent radicular symptoms (1/143) *ICBG:* 10.3% (14/136) - supplemental fixation for pseudoarthrosis (12/136) (all but 2 patients had radiographically solid fusion but repeat surgery done due to persistent pain) -supplemental fixation for persistent radicular symptoms (2/136)	-hematoma: 0.7% (1/136) Additional surgery due to complications: 0% (0/136) ***Hip pain (VAS scale 0-20*):*** *rhBMP2:* 0 at all time points *ICBG:* Discharge: 12.7 (134/136) 6 wks: 6.7 (132/136) 3 mos: 3.5 (134/136) 6 mos: 2.6 (132/136) 12 mos: 2.1 (130/136) 24 mos: 1.8 (117/136) *P* < .001 for all timepoints ***Patient very unhappy with appearance of graft site*** *ICBG:* Discharge: 9.7% (13/134) 6 wks: 3.7% (5/132) 3 mos: 2.2% (3/134) 6 mos: 3.7% (5/132) 12 mos: 3.8% (5/130) 24 mos: 2.6% (3/117)
Burkus (2003) (AHRQ ref 182) Integrated analysis, **includes all patients from Burkus 2003** **Lumbar spine** **On-label** Single-level primary anterior fusion with interbody fusion cages; performed via open or	NR	NR	<u>Revisions</u> *rhBMP2* Total: 0.4% (1/277) Open: 0% (0/143) Laproscopic: 0.7% (1/134) *Autograft* Total: 2.0% (8/402) Open: 0% (0/136) Laproscopic:	NR

laproscopic approach **rhBMP2 (n = 277) vs. ICBG (n = 402)**			3.0% (8/266) <u>Removals</u> *rhBMP2* Total: 1.4% (4/277) Open: 1.4% (2/143) Laproscopic: 1.5% (2/134) *Autograft* Total: 1.7% (7/402) Open: 0% (0/136) Laproscopic: 2.6% (7/266) <u>Supplemental fixations</u> *rhBMP2* Total: 6.1% (17/277) Open: 7.0% (10/143) Laproscopic: 5.2% (7/134) *Autograft* Total: 7.0% (28/402) Open: 10.3% (14/136) Laproscopic: 5.3% (14/266) <u>Reoperations</u> *rhBMP2* Total: 2.9% (8/277) Open: 4.2% (6/143) Laproscopic: 1.5% (2/134) *Autograft* Total: 8.0% (32/402) Open: 2.9% (4/136) Laproscopic: 10.5% (28/266) *P* = .004 for total reoperations for rhBMP2 vs. Autograft; P = ns for revisions removals and supplemental fixations	

FDA SSED: InFUSE (P000058) Integrated analysis **(**overlaps with Boden 2000[13], Burkus 2002[14], Burkus 2003[15]) **Lumbar spine** **On-label** rhBMP2 (n = 288) vs. ICBG (n = 139)	*rhBMP2 vs. ICBG* **Surgery results:** **Anatomical/technical difficulty:** 3.5% (10/288) (10 events) vs. 2.2% (3/139) (3 events) **Back and/or leg pain:** 0 vs. 0 events **Cancer**: 0% vs. 0% **Cardio/Vascular:** 2 vs. 0 events **Death:** 0 vs. 0 events **Dural injury:** 0 vs. 0 events **Gastrointestinal:** 1 vs. 0 events **Graft site related:** 0 vs. 0 events **Implant displacement/ loosening:** 0 vs. 0 events **Infection:** 0 vs. 0 events **Malpositioned implant:** 5 vs. 0 events **Neurological:** 0 vs. 0 events **Other:** 6 vs. 6 events **Other pain:** 0 vs. 0 events **Respiratory:** 0 vs. 0 events **Retrograde ejaculation:** 0 vs. 0 events **Spinal event:** 0 vs. 0 events **Subsidence:** 0 vs. 0 events	*rhBMP2 vs. ICBG* **Total # adverse events** (surgery to < 30 months) **Anatomical/technical difficulty:** 3.5% (10/288) (10 events) vs. 2.2% (3/139) (3 events) **Back and/or leg pain:** 22.6% (65/288) (72 events) vs. 21.6% (30/139) (33 events) **Cancer**: 0.3% (1/288) (1 event) vs. 0.7% (1/139) (1 event) **Cardio/Vascular:** 5.2% (15/288) (18 events) vs. 8.6% (12/139) (14 events) **Death:** 0% (0/288) vs. 0.7% (1/139) (pt had cardiovascular disease and died between 5-9 months postop). **Dural injury:** 0% (0/288) vs. 0.7% (1/139) (1 event) **Gastrointestinal:** 18.4% (53/288) (67 events) vs. 19.4% (27/139) (32 events) **Implant displacement/ loosening:** 1.7% (5/288) (5 events) vs. 0.7% (1/139) (1 event) **Infection:** 12.2% (35/288) (39 events) vs. 11.5% (16/139) (17 events) **Malpositioned implant:** 1.7% (5/288) (5 events) vs. 0% (0/139) (0 events) **Neurological:** 12.5% (36/288) (39 events) vs. 15.1% (21/139) (22 events) **Other:** 17.4% (50/288) (64 events) vs. 26.6% (37/139) (43 events) **Other pain:** 7.3% (21/288) (25 events) vs. 8.6% (12/139) (13 events)	*rhBMP2 vs. ICBG* **Non-union** (requiring second surgery)**:** 1.7% (5/288) (5 events) vs. 2.9% (4/139) (4 events) - Postop (1 day- 4 wks): 0 vs. 0 events -6 wks (4-9 wks): 0 vs. 0 events -3 mos (9 wks-5 mos): 1 vs. 0 events -6 mos (5-9 mos): 1 vs. 3 events -12 mos (9-19 mos): 2 vs. 0 events 24 mos: (19- < 30 mos): 1 vs. 1 events	*rhBMP2 vs. ICBG* **Graft site related adverse events:** 5.8% (8/139) (8 events) (details NR)

	Trauma: 0 vs. 0 events **Urogenital:** 1 vs. 0 events **Vascular intra-op:** 15 vs. 5 events **Vertebral fracture:** 0 vs. 0 events	**Respiratory:** 1.7% (5/288) (5 events) vs. 2.9% (4/139) (4 events) **Retrograde ejaculation:** 7.9% (11/140 males) (12 events) vs. 1.4% (1/70 males) (1 event): - Postop (1 day- 4 wks): 4 vs. 1 events -6 wks (4-9 wks): 5 vs. 0 events -3 mos (9 wks-5 mos): 1 vs. 0 events -6 mos (5-9 mos): 0 vs. 0 events -12 mos (9-19 mos): 2 vs. 0 events 24 mos: (19- < 30 mos): 0 vs. 0 events **Spinal event:** 8.3% (24/288) (27 events) vs. 11.5% (16/139) (17 events) **Subsidence:** 2.4% (7/288) (7 events) vs. 1.4% (2/139) (2 events) **Trauma:** 20.8% (60/288) (72 events) vs. 20.9% (29/139) (34 events) **Urogenital:** 11.5% (33/288) (37 events) vs. 7.2% (10/139) (11 events) **Vascular intra-op:** 4.9% (14/288) (15 events) vs. 3.6% (5/139) (5 events) **Vertebral fracture:** 0.3% (1/288) (1 event) vs. 0% (0/139) (0 events)		
Off-label use				
Boden (2002) (AHRQ ref 84) RCT **Lumbar spine Off-label** Single-level primary instrumented posterolateral lumbar fusion **rhBMP2 (n = 9) vs. rhBMP2/screw (n = 11)**	"There were no complications attributable to the rhBMP-2/BCP or TSRH internal fixation."	*rhBMP2 vs. rhBMP2/screw vs. ICBG/screw* ***Hematoma***: 22% (2/9) (epidural) vs. 9% (1/11) (required evacuation) vs. 0% (0/5) ***Persistent back pain:*** 11% (1/9) vs. 0 % (0/11) vs. 0% (0/5) ***Anti-BMP-2 antibodies:*** 4.5% (1/22) (BMP2 groups collapsed)(positive case	*rhBMP2* 11% (1/9) -revision at 8 months for persistent low back pain(1/9). *rhBMP2/screw* 18% (2/11) -decompression with resolution of leg pain (1/11) -revision at 1 year (1/11)	***Hip pain (VAS scale 0-20*):*** *rhBMP2:* NR *rhBMP2/Screw:* NR *ICBG/Screw:* Discharge: 16.0 (± 0.7 SEM) 17 mos (mean): 5.2 (±2.3 SEM)

vs. ICBG/screw (n = 5)		was transient upon subsequent testing) vs. 0% (0/4)	*ICBG/Screw* 0% (0/5)	At 17 mos follow-up, mean not different from zero (*P* = .088)
Burkus (2005) (AHRQ ref 85) **Burkus (2006)** (excluded by AHRQ; safety data reported here) RCT **Lumbar Spine** **Off-label** Primary single-level anterior lumbar fusion with a pair of threaded allograft cortical bone dowels (CBD) **rhBMP2 (n=79) vs. ICBG (n=52)**	*rhBMP2 vs. ICBG* ***infection* (Burkus 2006):** 0% vs. 0%	*rhBMP2 vs. ICBG* ***heterotopic bone formation (bone remodeling):*** 18% (14/79) vs. 0% (0/52) (transient localized areas of bone remodeling in the vertebral body adjacent to an allograft dowel. All resolved by 24 mos) ***(Burkus 2006):*** -Not influenced by fusion level (*P* = .2145) -All zones filled with new trabecular bone formation at 24 mos. -no association with development of bone remodeling zones -no evidence of radiolucencies at 12 mos. after surgery. ***Graft migration:*** 0% (0/79) vs. 0% (0/52) ***Graft extrusion:*** 0% (0/79) vs. 0% (0/52) ***Implant fracture:*** 0% (0/79) vs. 0% (0/52) ***elevated anti-rhBMP2 antibodies:*** 0% (0/78) vs. 0% (0/49) ***elevated anti-bovine collagen antibodies:*** 9% (7/78) vs. 8% (4/49) ***Allograft incorporation (Burkus 2006):*** ***-complete incorporation (healing):*** 6 mos.: 72% vs. 45% 12 mos.: 96% vs. 66% 24 mos.: 100% vs. 79% ***-partial incorporation (healing):*** 6 mos.: 27% vs. 38% 12 mos.: 4% vs. 23% 24 mos.: 0% vs 10% ***-no incorporation***	*rhBMP2 vs. ICBG* ***Supplemental fixation:*** 3% (2/79) vs. 15% (8/52) No other additional procedures were performed ***Perioperative disc material removal (Burkus 2006):*** 0 vs. 1 (early postop; no interval given) ***Supplemental fixation (Burkus 2006):*** 1 vs 5 (>24 mos. postop)	"Pain at the donor site was similar to previous reports but the pain was observed to persist at a slightly higher rate of 46.5%."

		(healing): 6 mos.:1% vs. 17% 12 mos.: 0% vs. 11% 24 mos.: 0% vs. 11% "At no follow-up interval was new bone formation found to extend outside of the disc space in either the investigational or the control group" (Burkus 2006) -***elevated anti-bovine collagen antibodies in patients with vs. without bone remodeling zones***: 14% (2/14) vs. 18% (12/65) -***elevated anti-rhBMP-2 antibodies in patients with vs. without bone remodeling zones***: 0% vs. 0% ***infection* (Burkus 2006):** 0% vs. 0%		
Dawson (2009) (AHRQ ref 86) RCT **Lumbar spine** **Off-label** Single-level primary instrumented posterolateral lumbar fusion **rhBMP2 (n = 25) vs. ICBG (n = 21)**	***Malpositioned screws:*** 4% (1/25) (at 1 day; see second surgeries) vs. 0% (0/21) ***Durotomy:*** 4% (1/25) vs. 5% (1/21) ***Wound infection (surgical site)***: 4% (1/25) vs. 5% (1/21) (resolved with antibiotics)	*rhBMP2 vs. ICBG* ***Death (cause NR):*** 4% (1/25) vs. 0% (0/21)	*rhBMP2 vs. ICBG* 8% (2/25) (revision at 1 day due to malpositioned screws in one pt; removal of hardware at 6 mos in other patient) vs. 10% (2/21) (revision for psuedoarthrosis between 12-24 mos f/u)	***Graft site infection:*** 5% (1/21)
Dimar (2009) (AHRQ ref 86) RCT **Lumbar spine** **Off-label** Single-level primary instrumented posterolateral lumbar fusion **rhBMP2 (n = 239) vs. ICBG (n = 224)**	*rhBMP2 vs. ICBG* **Any adverse event (operative):** 8.4% (20/239) vs. 8.9% (20/224) **Any possible implant or implant-related event** (noted where appropriate below)**:** 0% (0/239) vs. 1.3% (3/224) ***Anatomic/technical***	*rhBMP2 vs. ICBG* **Any adverse event (f/u up to 24 mos.):** 87.4% (200/239) vs. 88.4% (198/224) (*P* = .777) **Any possible implant or implant-related event (f/u up to 24 mos.)** (noted where appropriate below)**:** 8.8% (21/239) vs. 15.6% (35/224) (*P* = .032) ***Anatomic/technical difficulty* event (f/u up to**	*rhBMP2 vs. ICBG* **Any:** 8.4% (20/239) vs. 16.1% (36/224) (*P* = .015) **Revision:** 1.7% (4/239) vs. 1.8% (4/224) (*P* = NR) **Nonelective removal** (due to nonunion OR adverse event OR not at discretion of	*rhBMP2 vs. ICBG* **Any** (details NR) 0% (0/239) vs. 7.6% (17/224) (*P* < .001) *See infection under Adverse Events* ***Graft site pain (VAS 0-20*)*** *(ICBG group only):* Discharge: 11.3 1.5 mos: 7.9 3 mos: 6.3 24 mos: 5.1

NOTE. Contains patients in Glassman (2007), Dimar (2006), and Glassman (2005)	***difficulty:*** 0.4% (1/239) vs. 0% (0/224) **Arthritis/bursitis** 0% (0/239) vs. 0% (0/224) (0 vs. 0 possibly implant-related) **Back and/or leg pain:** 0% (0/239) vs. 0% (0/224) (0 vs. 0 possibly implant-related) ***Cardiovascular (details NR):*** 0.8% (2/239) vs. 0% (0/224) **Carpal tunnel syndrome:** 0% (0/239) vs. 0% (0/224) **Death**: 0% (0/239) vs. 0% (0/224) ***Dural injury:*** 5.4% (13/239) vs 8.0% (18/224) (0 vs. 1 possibly implant-related) **Gastrointestinal:** 0% (0/239) vs. 0% (0/224) (0 vs. 0 possibly implant related) ***Malpositioned implant:*** 0% (0/239) vs. 0.4% (1/224) ***Implant displacement and/or loosening:*** 0% (0/239) vs. 0.4% (1/224) (0 vs. 1 possibly implant-related) **Infection** (details NR)**:** 0% (0/239) vs. 0% (0/224) **Neurological** (details NR)**:** 0% (0/239) vs. 0% (0/224) (0 vs. 0 possibly implant-related)	**24 mos.)*:*** 0.4% (1/239) vs. 0% (0/224) **Arthritis/bursitis** 9.2% (22/239) vs. 7.6% (17/224) (*P* = .616) (0 vs 2 possibly implant-related; *P* = .234) **Back and/or leg pain:** 43.5% (104/239) vs. 40.2% (90/224) (*P* = .510) (4 vs. 5 possibly implant-related; *P* = .745) **Cancer:** 3.3% (8/239) (basal cell carcinoma, lung, lymphoma, ovarian, pancreatic, prostate, squamous cell carcinoma, vocal cord) vs. 0.9% (2/224) (colon, lymphoma) (*P* = .107) ***Cardiovascular** (details NR):* 21.8% (52/239) vs. 24.1% (54/224) (*P* = .581) **Carpal tunnel syndrome:** 3.8% (9/239) vs. 2.7% (6/224) (*P* = .604) **Death** ("causes unrelated to surgery")**:** 1.3% (3/239) vs. 1.8% (4/224) (*P* = .717) **Dural injury:** 5.9% (14/239) vs. 8.0% (18/224) (*P* = .367) (0 vs. 1 possibly implant-related; *P* = .484) **Gastrointestinal:** 15.5% (37/239) vs. 14.7% (33/224) (*P* = .897) **Heterotopic ossification in surrounding tissue:** 0% (0/239) vs. 0% (0/224) **Implant displacement and/or loosening:** 0.4% (1/239) vs. 1.3% (3/224) (*P* = .358) (1 vs. 3 possibly implant-related; *P* = .358) **Infection** (details NR)**:** 16.3% (39/239) vs. 20.1% (45/224) (*P* = .335)	patient or investigator)**:** 4.2% (10/239) vs. 10.3% (23/224) (*P* = NR) **Supplemental fixation:** 2.5% (6/239) vs. 4.0% (9/224) (*P* = NR) **Nonunion failure** (patients who required additional surgery due to nonunion)**:** 2.5% (6/239) vs. 8.0% (18/224) (*P* = .011) (6 vs. 18 possibly implant-related; *P* = .011) **Nonunion outcome pending** (description NR)**:** 2.1% (5/239) vs. 2.2% (5/224) (*P* = .011) (5 vs. 4 possibly implant-related; *P* = 1.000)	***Graft site pain (% of patients experiencing pain):*** 24 mos: 60% (108/180 reporting)

	Other** (not specified):* 0.4% (1/239) vs. 0% (0/224) ***Other pain** (details NR):* 0% (0/239) vs. 0% (0/224) ***Respiratory** (details NR):* 0% (0/239) vs. 0% (0/224) ***Spondylosis or stenosis (any level): 0% (0/239) vs. 0% (0/224) ***Trauma** (details NR):* 0% (0/239) vs. 0% (0/224) ***Urogenital** (details NR):* 0% (0/239) vs. 0% (0/224) ***Vertebral fracture:*** 1.3% (3/239) vs. 1.3% (3/224) (0 vs. 1 possibly implant-related)	***Malpositioned implant:*** 2.1% (5/239) vs. 0.9% (2/224) (*P* = .451) (4 vs. 2 possibly implant-related; *P* = .686) **Neurological** (details NR, neurological outcomes not reported otherwise)**:** 29.3% (70/239) vs. 26.8% (60/224) (*P* = .605) (2 vs. 1 possibly implant-related; *P* = 1.000) ***Other** (details NR):* 29.3% (70/239) vs. 27.7% (62/224) (*P* = .758) ***Other pain** (details NR):* 12.1% (29/239) vs. 12.5% (28/224) (*P* = 1.000) ***Respiratory** (details NR):* 6.3% (15/239) vs. 5.4% (12/224) (*P* = .697) ***Spondylosis or stenosis (any level):*** 7.1% (17/239) vs. 8.0% (18/224) (*P* = .728) ***Trauma** (details NR):* 28.0% (67/239) vs. 26.3% (59/224) (*P* = .754) ***Urogenital** (details NR):* 10.9% (26/239) vs. 9.4% (21/224) (*P* = .646) ***Vertebral fracture:*** 1.3% (3/239) vs. 2.2% (5/224) (*P* = .492) (0 vs. 1 possibly implant-related; *P* = .484)		
FDA Executive Summary: AMPLIFY (P050036) **Lumbar spine IDE study** rhBMP2 (n = 239) vs. ICBG (n = 224) at 24 months rhBMP2 (n = 222) vs. ICBG (n = 210) at 60 months	*rhBMP2 vs. ICBG* **Surgery results:** **Anatomical/technical difficulty:** 0.4% (1/239) (1 event) vs. NR **Arthritis/bursitis:** NR **Back and/or leg pain:** NR	*rhBMP2 vs. ICBG* **Total # adverse events** (surgery to ≤ 24 months **|** surgery to ≤ 60 months) **Anatomical/technical difficulty:** 0.4% (1/239) (1 event) vs. 0% (0/224) (0 events) **|** 0.4% (1/222) (1 events) vs. 0% (0/210) (0 events) **Arthritis/bursitis:** 9.6%	*rhBMP2 vs. ICBG* **Non-union:** 4.2% (10/239) (10 events) vs. 10.3% (23/224) (23 events) at 24 months **|** 4.6% (11/222) (11 events) vs. 11.2% (25/210) (25 events) at 60 months - Postop (1 day- 4	*rhBMP2 vs. ICBG* **Graft site related adverse events:** 7.6% (17/224) (17 events) (details NR)

	Cancer: NR **Cardio/Vascular:** 0.9% (2/239) (2 events) vs. NR **Carpal tunnel syndrome:** NR **Death:** NR **Dural injury:** 5.4% (13/239) (13 events) vs 8.0% (18/224) (18 events) **Gastrointestinal:** NR **Graft site related:** NR **Implant displacement/ loosening:** NR **Infection:** NR **Malpositioned implant:** 0.4% (1/239) (1 event) vs. NR **Neurological:** NR **Other:** 0.4% (1/239) (1 event) vs. NR **Other pain:** NR **Respiratory:** NR **Spinal event:** NR **Trauma:** NR **Urogenital:** 0 events vs. NR **Vertebral fracture:** 1.3% (3/239) (3 events) vs. 1.3% (3/224) (3 events)	(23/239) (24 events) vs. 7.6% (17/224) (19 events) **|** 13.0% (31/222) (37 events) vs. 12.1% (27/210) (34 events) **Back and/or leg pain:** 43.9% (105/239) (139 events) vs. 39.7% (89/224) (110 events) **|** 54.8% (131/222) (219 events) vs. 55.4% (124/210) (190 events) **Cancer**: 3.8% (9/239) (9 events) vs. 0.9% (2/224) (2 events) (*P* = NS) **|** 5.0% (12/222) (15 events) vs. 2.1% (5/210) (5 events) (*P* borderline) **Cardio/Vascular:** 22.2% (53/239) (72 events) vs. 24.1% (54/224) (67 events) **|** 30.5% (73/222) (101 events) vs. 28.1% (63/210) (84 events) **Carpal tunnel syndrome:** 3.8% (9/239) (9 events) vs. 2.7% (6/224) (6 events) **|** 3.8% (9/222) (9 events) vs. 3.6% (8/210) (8 events) **Death:** 1.3% (3/239) (3 events) vs. 1.8% (4/224) (4 events) **|** 2.5% (6/222) (6 events) vs. 3.1% (7/210) (7 events) **Dural injury:** 5.9% (14/239) (14 events) vs. 8.0% (18/224) (18 events) **|** 5.9% (14/222) (14 events) vs. 8.5% (19/210) (20 events) **Gastrointestinal:** 15.5% (37/239) (43 events) vs. 14.7% (33/224) (43 events) **|** 24.3% (58/222) (75 events) vs.22.8 % (51/210) (70 events) **Graft site related:** NR) **|** 0% (0/222) (0 events) vs.8.5 % (19/210) (19 events) **Implant displacement/ loosening:** 0.4% (1/239)	wks): NR -6 wks (4-9 wks): NR -3 mos (9 wks-5 mos): 1 vs. 8 events -6 mos (5-9 mos): NR vs. 6 events -12 mos (9-19 mos): 8 vs. 6 events 24 mos: (19- < 30 mos): 1 vs. 3 events **All second surgeries:** 46.0% (110/239) vs. 62.5% (140/224) at 24 months -Revisions 1.7% (4/239) vs. 1.8% (4/224) -Removals total 5.4% (13/239) vs. 12.5% (28/224) -Removals non-elective 4.2% (10/239) vs. 9.8% (22/224) -Removals elective 1.3% (3/239) vs. 2.7% (6/224) -Supplemental fixations 2.5% (6/239) vs. 4.0% (9/224) -Reoperations 5.0% (12/239) vs. 4.9% (11/224) -Other 25.9% (62/239) vs. 26.8% (60/224)	

		(1 event) vs. 0.9% (2/224) (2 events) | 0.4% (1/222) (1 event) vs. 0.9% (2/210) (2 events) **Infection:** 16.3% (39/239) (52 events) vs. 20.1% (45/224) (51 events) | 18.8% (45/222) (60 events) vs. 22.8% (51/210) (64 events) **Malpositioned implant:** 2.1% (5/239) (5 events) vs. 0.9% (2/224) (2 events) | 2.1% (5/222) (5 events) vs. 0.9% (2/210) (2 events) **Neurological:** 29.3% (70/239) (85 events) vs. 26.8% (60/224) (74 events) | 35.6% (85/222) (113 events) vs. 32.1% (72/210) (98 events) **Other:** 29.3% (70/239) (101 events) vs. 27.7% (62/224) (91 events) | 37.2% (89/222) (174 events) vs. 35.7% (80/210) (147 events) **Other pain:** 12.1% (29/239) (31 events) vs. 12.9% (29/224) (32 events) | 19.7% (47/222) (58 events) vs. 20.1% (45/210) (59 events) **Respiratory:** 6.7% (16/239) (17 events) vs. 5.4% (12/224) (13 events) | 6.7% (16/222) (17 events) vs. 6.3% (14/210) (18 events) **Spinal event - all:** 7.1% (17/239) (18 events) vs. 8.5% (19/224) (22 events) | 11.7% (28/222) (30 events) vs. 9.8% (22/210) (26 events) **Spinal event - cervical:** NR | 6.3% (15/222) (16 events) vs. 6.3% (14/210) (15 events) **Spinal event - lumbar:** NR | 5.4% (13/222) (13 events) vs. 4.5% (10/210) (10		

		events) **Spinal event - thoracic:** NR **|** 0.4% (1/222) (1 event) vs. 0.4% (1/210) (1 event) **Trauma:** 28.9% (69/239) (91 events) vs. 26.3% (59/224) (70 events) **|** 38.5% (92/222) (131 events) vs. 33.9% (76/210) (104 events) **Urogenital:** 11.3% (27/239) (28 events) vs. 9.4% (21/224) (24 events) **|** 13.8% (33/222) (37 events) vs. 12.5% (28/210) (32 events) **Vertebral fracture:** 1.3% (3/239) (3 events) vs. 1.8% (4/224) (4 events) **|** 1.3% (3/222) (3 events) vs. 1.8% (4/210) (4 events)		
Glassman (2008) (AHRQ ref 87) RCT **Lumbar spine Off-label** Single- or multi-level primary instrumented posterolateral lumbar fusion plus rhBMP2 or ICBG **rhBMP2 (n=50) vs. ICBG (n=52)**	"None of the complications were directly attributable to either the ICBG harvest or the rhBMP2 use) *rhBMP2 vs. ICBG* **Perioperative complications** (up to 3 months) (any): 16% (8/50) vs. 23% (12/52) (number of patients having complications) **Cardiac** (details NR)**:** 2% (1/50) vs. 13% (7/52) **Wound infection:** 2% (1/50) vs. 8% (4/52) (see also second surgeries) **Back or leg pain:** 0% (0/50) vs. 6% (3/52) (requiring readmission or epidural steroid treatment) **Gastrointestinal:** 4% (2/50) vs. 6%	*rhBMP2 vs. ICBG* **Progressive radiculopathy:** 0% (0/50) vs. 2% (1/52)	*rhBMP2 vs. ICBG* **All second surgeries:** 8% (4/50) vs. 21% (11/52) **- Treatment of wound infection (details NR):** 2% (1/50) vs. 4% (2/52) - **Repositioning of pedicle screw:** 0% (0/50) vs. 2% (1/52) - **Extension of fusion for adjacent level compression fracture:** 2% (1/50) vs. 0% (0/52) - **Revision for nonunion:** 2% (1/50) vs. 10% (5/52) - **Late screw removal:** 0% (0/50) vs. 2% (1/52) (due to	

	(3/52) **Urinary tract infection:** 2% (1/50) vs. 2% (1/52) **Neurological deficit:** 0% (0/50) vs. 2% (1/52) **Line-related sepsis:** 2% (1/50) vs. 0% (0/52) **Broken toe:** 2% (1/50) vs. 0% (0/52) **Shingles:** 2% (1/50) vs. 0% (0/52) **Multiple complications:** 0 vs. 6 (patients) **Total number of perioperative complications:** 8 vs. 20 (*P* = .014)		progressive radiculopathy and weakness, refused exploration or revision of fusion) - **Pain pump insertion** 0% (0/50) vs. 2% (1/52) - **Revision for adjacent level degeneration:** 2% (1/50) vs. 2% (1/52)	
Haid (2004) (AHRQ ref 88) RCT **Lumbar spine** **Off-label** Single-level primary posterior lumbar interbody fusion (PLIF) with interbody fusion cages **rhBMP2 (n = 34)** **vs. ICBG (n = 33)**	"No unanticipated device-related [surgical] adverse events occurred in either treatment group." *rhBMP2 vs. ICBG* Deep vein thrombosis: 0% (0/34) vs. 3% (1/33) (treated with anticoagulation medications) Neurological complications: 41% (14/34) (16 events in 14 patients) vs. 42% (14/33) (18 events) Dural tears: 9% (3/34) vs. 6% (2/33)	*rhBMP2 vs. ICBG* ***Graft subsidence:*** 6% (2/34) vs. 6% (2/33) (cages countersunk 3mm or more from the posterior margin) ***Spondylolisthesis:*** -Any (new or residual): 12% (8/67) (group NR) -New: 3% (2/67) (group NR) ***Extradiscal bone formation*** *(outside disc space and into the spinal canal or neuroforamina):* 75% (24/32) vs. 13% (4/31) (*P* <.0001) (scans or radiographs unavailable in 4 patients) -strongly associated with cage placement within 2 mm of the margin of the posterior vertebral cortex: 77% (23/30) vs 12% (# pts	*rhBMP2* 18% (6/34) (any secondary spinal surgical procedure) -Failures: 9% (3/34) (revision surgery at the same level; not radiographic fusion failures) -Fusion at a different spinal level: 9% (3/34) *ICBG* 18% (6/33) (any secondary spinal surgical procedure) -Failures: 9% (3/33) - Fusion at a different spinal level: 9% (3/33)	***ICBG site complications:*** 6% (2/33) (1 case of pain and 1 case of hematoma at the graft site; neither required surgery) ***Hip pain (VAS scale 0-20*):*** *rhBMP2:* NR *ICBG:* Discharge: 11.6 24 mos: 5.5 (60% still experienced pain at the graft site (i.e. had scores >0); 13% of patients stated the appearance of the graft site bothered them some)

		with this type cage placement NR for ICBG gp) -no correlation with leg pain (7/22 ICBG patients with increased leg pain had no bone formation outside the disc space) anti-rhBMP2 antibodies: 0% (0/34) vs. 0% (0/33) **anti-human Type 1 collagen antibodies**: 0% (0/34) vs. 0% (0/33) **anti-bovine Type 1 collagen antibodies:** 9% (3/34) vs. 15% (5/33) (positive antibody detection at 3 times baseline) (no clinical sequelae)		
Glassman (2007) (AHRQ ref 99) Retrospective cohort with historical control **Lumbar spine** **Off-label** Single- or multilevel primary or revision instrumented posterolateral fusion **rhBMP2 (n = 91) vs. ICBG (n = 35)**	NR	NR	*rhBMP2 vs. ICBG:* Reexploration in patients initially enrolled for revision surgery ONLY (details NR): 31% (5/16) vs. NR	NR
Mummaneni (2004) (AHRQ ref 100) Retrospective cohort study **Lumbar spine** **Off-label** Single- or multilevel primary instrumented transforaminal lumbar interbody fusion (TLIF) with interbody fusion cages **rhBMP2/AGB (± ICBG) (n = 25)** **vs. ICBG (n = 19)**	*rhBMP2/AGB ± ICBG vs. ICBG* **CSF leak**: 10% (2/21) vs. 11% (2/19) **Paresis (L-5):** 5% (1/21) vs. 5% (1/19)	*rhBMP2/AGB ± ICBG vs. ICBG* **Worsening of preoperative partial foot drop:** 0% (0/21) vs. 5% (1/19) **Weakness of ankle dorsiflexion:** 5% (1/21) vs. 0% (0/19) (at 9 mos. follow-up, problem resolved) **Pseudarthrosis:** 0% (0/20) vs. 5% (1/19) **Foraminal bone formation:** 0% (0/21) vs NR	NR	58% of patients complained of donor site pain 6 mos. after surgery (mean pain grade of 5 out of 10 VAS) (this group includes ICBG group in addition to rhBMP2/AGB+ICBG)
Pradhan 2006 (AHRQ ref 101)	NR	"In the cases of nonunion	*rhBMP2/ACS vs. ICBG*	NR

Prospective cohort study (historical control)† **Lumbar spine** **Off-label** Single level primary anterior lumbar interbody fusion (ALIF) with femoral ring allograft (FRA) **rhBMP2/ACS (n = 9) vs. ICBG (n = 27)**		with BMP, extensive osteolysis of and around the FRA was seen, causing, fracture, fragmentation, and collapse of the graft." "In the ICBG group the FRA never seemed to be completely resorbed."	All second surgeries: 33% (3/9) vs. 26% (7/27) (all second surgeries were salvage posterior fusions to treat nonunion)	
Singh 2006 (AHRQ ref 102) Prospective case-control study **Lumbar spine** **Off-label** Single- or multi-level primary instrumented posterolateral fusion **rhBMP2/ICBG (n = 39) vs. ICBG/local autograft (n =11)**	*rhBMP2/ICBG vs. ICBG/local autograft* Dural tear: 5% (2/39) vs. NR	*rhBMP2/ICBG vs. ICBG/local autograft* **Ectopic muscle ossification**: 0% (0/39) vs. NR **Intra- or extradural ossification**: 0% (0/39) vs. NR **Laminar bone regrowth**: 0% (0/39) vs. NR	*rhBMP2/ICBG vs. ICBG/local autograft* **Decompression**: 3% (1/39) vs. NR (trasitional stenosis above fusion mass; at 10 mos. follow-up; this patient showed 'degenerative facet changes' preoperatively).	NR
Slosar 2007 (AHRQ ref 103) Prospective cohort study **Lumbar spine** **Off-label** Single- or multi-level primary instrumented anterior lumbar interbody fusion (ALIF) with femoral ring allograft (FRA) **rhBMP2/ACS (n = 45) vs. ALG (allograft bone chips) (n = 30)**	*rhBMP2 vs. ALG* **Dural tear** (rent): 2% (1/45) vs. 0% (0/30) **Deep (posterior) wound infection**: 2% (1/45) (required irrigation, debridement, delayed closure and intravenous antibiotics) vs. 0% (0/30) **Superficial (anterior) wound dehiscence**: 0% (0/45) vs. 3% (1/30)	"There were no complications attributable to the use of rhBMP-2." rhBMP2 vs. ALG **Ectopic bone formation:** 0% (0/45) vs. 0% (0/30) **Osteolysis of allograft:** 0% (0/45) vs. 0% (0/30) **Fragmentation of allograft:** 0% (0/45) vs. 0% (0/30) **Pseudarthrosis:** 0% (0/45) vs. 17% (5/30) (4/5 patients received salvage posterolateral fusion; the last patient is pending)	*rhBMP2 vs. ALG* **Treatment of deep wound infection** (see perioperative complications)**:** 2% (1/45) vs. 0% (0/30) **Salvage posterolateral fusion** (for pseudarthrosis)**:** 0% (0/45) vs. 13% (4/30) (with 1/30 pending)	n/a
Carragee (2011) **Retrospective cohort** **Lumbar spine**	NR	*rhBMP2 vs. ICBG* ***RE:*** 7.2% (5/69) vs. 0.6% (1/174) *P* = .003	NR	NR

		3/5 (60.0%) with RE in rhBMP2 group had early osteolysis and 1/5 (20.0%) had extensive osteolysis with fracture of the sacral body seen on plain radiograph in early postop period ***RE, 1-level L5/S1 fusion:*** 6.7% (3/45) vs. 0% (0/110) *P* = .023 ***RE, 2-level L4/L5 and L5/S1 fusion:*** 8.3% (2/24) vs. 1.6% (1/64) *P* = .179 ***Resolution of RE 1 year postop:*** 40.0% (2/5) vs. 100% (1/1)		
Crawford et al (2010) **Retrospective cohort with historical control** **Sacrum** (appears to contain the same patients reported in Maeda (2009))	*rhBMP2 vs. autograft* ***Total surgical complications:*** 5 versus 4 complications (not patients) *P* = .181 ***Nerve root deficit (all resolved after reoperation):*** 5.6% (2/36) vs. 4.2% (1/24) ***Vein tear with ASF:*** 0% (0/36) vs. 8.3% (2/24) ***ASF aborted due to scared down interior vena cava (IVC):*** 2.8% (1/36); autograft n/a ***Deep wound infection:*** 2.8% (1/36) vs. 4.2% (1/24) ***Ulnar nerve paresthesia:*** 2.8% (1/36) vs. 0% (0/24)	*rhBMP2 vs. autograft* ***Total follow-up complications:*** 9 vs. 12 complications (not patients); *P* = .058 ***Total medical complications*** *(appendicitis, UTI, pneumonia):* 4 versus 1 complications (not patients); *P* = .639 ***Appendicitis 3 mo. postop:*** 0% (0/36) vs. 4.2% (1/24) ***Urinary tract infection:*** 8.3% (3/36) vs. 0% (0/24) ***Pneumonia (readmit):*** 2.8% (1/36) vs. 0% (0/24) -- ***Broken rod between S1 and Iliac:*** 2.8% (1/36) vs. 8.3% (2/24) ***Broken rod L5-S1:*** 2.8% (1/36) vs. 16.7% (4/24) ***Broken rod L4-L5:*** 5.6% (2/36) vs. 4.2% (1/24) ***Broken rod L3-L4:*** 0% (0/36) vs. 4.2% (1/24)	*rhBMP2 vs. autograft* ***Reoperation for pseudarthosis:*** 5.6% (2/36) vs. 12.5% (3/24); *P* = .380 ***Iliac screw removed:*** 8.3% (3/36) vs. 8.3% (2/24) ***Reoperation for nerve root deficit:*** 5.6% (2/36) vs. 4.2% (1/24)	n/a

		Broken rod L2-L3: 2.8% (1/36) vs. 0% (0/24) ***Vertebral compression fractures T8-T9:*** 2.8% (1/36) vs. 0% (0/24) ***Coronal imbalance:*** 0% (0/36) vs. 8.3% (2/24) ***Pseudarthrosis:*** 11.1% (4/36) vs. 20.8% (5/24) ***Tissue swelling:*** 0% (0/36) vs. 0% (0/24) ***Hematoma:*** 0% (0/36) vs. 0% (0/24) ***Seroma:*** 0% (0/36) vs. 0% (0/24) ***Heterotopic ossification:*** 0% (0/36) vs. 0% (0/24) ***Delayed radiculopathy:*** 0% (0/36) vs. 0% (0/24)		
Howard et al. (2011) **Cross-sectional** **Lumbar spine**	NR	NR	NR	***Iliac graft site pain score (mean, 0–10):*** rhBMP2: 50.8% (30/59) ICBG: 56.6% (30/53) *P* = ns **Severity of pain on palpation (mean ± SD):** rhBMP2: 3.6 ± 3.8 ICBG: 3.8 ± 3.2 *P* = ns
Joseph et al. (2007) **Prospective cohort** **Lumbar spine**	NR	*rhBMP2 vs. local autograft* ***Nonunion*** 6 mos: 9% (2/23) vs. 50% (5/10) (*P* = .016) 12 mos: 0% (0/23) vs. 10% (1/10) (pt has tolerable mechnical LBP with heavy labor) ***Heterotopic (extradiscal) bone formation*** 20.8% (5/24) vs. 8.3%	NR	n/a

		(1/12) levels (*P* = .64) (no clinical sequelae) ***Ectopic bone formation*** 0% (0/24) vs. 0% (0/12) levels ***Paraspinal bone formation*** 0% (0/24) vs. 0% (0/12) levels		
Latzman et al (2010) **Retrospective cohort** **Lumbar or lumbosacral spine**	NR	***Renal function:*** **BUN (mg/dL; mean ± SD)** *Preop* rhBMP2: 13.7 ± 4.1 auto/allograft: 15.4 ± 5.2 *Postop* rhBMP2: 19.7 ± 15.7 auto/allograft: 17.2 ± 7.1 *P* = ns **Creatinine (mg/dL; mean ± SD)** *Preop* rhBMP2: 0.8 ± 0.3 auto/allograft: 1.0 ± 0.2 *Postop* rhBMP2: 1.1 ± 0.9 auto/allograft: 1.1 ± 0.3 *P* = ns Note increased SD among rhBMP2 patients ***Transient renal insufficiency (BUN > 30 mg/dL; creatinine > 1.5 mg/dL):*** rhBMP: 12.5% (3/24) auto/allograft: 0% (0/105) *P* = .006 No patient experienced progressive renal failure and all had returned to preop values by 2 months after surgery No sepsis or other infections or wound breakdown 2/3 patients with renal insufficiency experienced transient supraventricular tachycardia, mental status changes, and fever – both had received 16 cc of rbBMP2 rather than 8 cc in 2 of their 3 postoperative courses	NR	NR

		These 2 patients were both diagnosed with malignancies during the 8 months after surgery ***New cancer diagnoses*** rhBMP2: 16.7% (4/24) auto/allograft: 7.6% (8/105) *P* = ns		
Lee et al. (2010) **Retrospective cohort** **Lumbar spine**	***rhBMP2 age ≥ 65 years vs. ICBG age ≥ 65 years*** **Total**: 32.4% (11/34) vs. 48.8% (20/41) **Dural tear**: 2.9% (1/34) vs. 7.3% (3/41) **Cardiac problems** (details NR): 5.9% (2/34) vs. 9.8% (4/41) **GI problems**: 5.9% (2/34) vs. 9.8% (4/41) **UTI**: 2.9% (1/34) vs. 4.9% (2/41) **Neurological deficit**: 2.9% (1/34) vs. 2.4% (1/41) **DVT**: 8.8% (3/34) vs. 12.2.% (5/41) **Wound infection**: 2.9% (1/34) vs. 2.4% (1/41) *P* = ns for all comparisons	NR	**Revision surgery** rhBMP2 age ≥ 65 years: 16.7% (1/6) rhBMP2 age < 65 years: 0% (0/3) ICBG age age ≥ 65 years: 22.2% (2/9) *P* = ns	NR
Rihn et al. (2009) **Retrospective cohort** **Lumbar spine**	***Malpositioned instrumentation*** rhBMP2: 2.3% (2/86) ICBG: 0% (0/33)	***Any complication*** *(includes malpositioned instrumentation & donor site infection or pain)* rhBMP2: 29.1% (25/86) ICBG: 45.5% (15/86) (*P* = .09) ***Total number of complications:*** rhBMP2: 37 ICBG: 18 ***Lumbar infection*** rhBMP2: 3.5% (3/86) ICBG: 6.1% (2/33) (*P* = NR) ***Lumbar hematoma*** rhBMP2: 1.2% (1/86) ICBG: 3.0% (1/33) (*P* = NR) ***Lumbar seroma*** rhBMP2: 1.2% (1/86) ICBG: 0% (0/33) (*P* = NR)	***Reoperation*** rhBMP2: 9.3 (8/86) ICBG: 12.1% (4/33) (*P* = .65) **Reasons for reoperation** **Reasons for reoperation** **<u>BMP</u>** Retained drain (n = 1) Lumbar hematoma (n = 1) Lumbar seroma (n = 1) Malpositioned screw with radiculitis (n = 1) Ectopic bone formation within neuraforaimen with postop radiculitis (n = 1) Lumbar wound	***Persistent donor-site pain*** rhBMP2: n/a ICBG: 30.3% (10/33) ***Reoperation*** rhBMP2: n/a ICBG: 3.0% (1/33) (due to donor-site infection) ***Donor-site infection*** rhBMP2: n/a ICBG: 3.0% (1/33) (required reoperation)

		Radiculitis rhBMP2: 14.0% (12/86) ICBG: 3.0% (1/33) (*P* = .08) ***Ectopic bone formation*** rhBMP2: 2.3% (2/86) ICBG: 0% (0/33) (*P* = NR) ***Vertebral osteolysis*** rhBMP2: 5.8% (5/86) ICBG: 0% (0/33) (*P* = NR) ***Dural tear*** rhBMP2: 4.7% (4/86) ICBG: 0% (0/33) (*P* = NR) ***Nonunion*** rhBMP2: 3.5% (3/86) ICBG: 3.0% (1/33) (*P* = NR) ***Urinary tract infection*** rhBMP2: 2.3% (2/86) ICBG: 3.0% (1/33) (*P* = NR) ***Ileus*** rhBMP2: 1.2% (1/86) ICBG: 3.0% (1/33) (*P* = NR) ***Retained drain*** rhBMP2: 1.2% (1/86) ICBG: 0% (0/33) (*P* = NR)	infection (n = 3) <u>**ICBG**</u> Lumbar hematoma (n = 1) Lumbar wound infection (n = 2) ICBG donor site infection (n = 1)	
Taghavi et al. (2010) **Retrospective cohort** **Lumbar spine**	***Dural tear*** rhBMP2: 4.2% (1/24) BMAA: 0% (0/18) Autograft: 5.0% (1/20)	***Psuedarthrosis*** rhBMP2: 0% (0/24) BMAA: 22.2% (4/18) Autograft: 0% (0/20)	***Hardware removal due to persistent irritation*** rhBMP1: 8.3% (2/24) BMAA: 5.6% (1/18) Autograft: 10.0% (2/20) ***Revision*** rhBMP2: 0% (0/24) BMAA: 16.7% (3/18)	***Persistent donor-site pain*** BMAA: 0% (0/18) Autograft: 20.0% (4/20)

			Autograft: 0% (0/20)	
Vaidya, Weir et al. (2007) [37 37 36 35] **Prospective cohort** **Lumbar (+ cervical, NR here) spine**	NR	*rhBMP2/allograft vs. DBM/allograft* ***Nonunion (lumbar only)*** 0% (0/25) vs 0% (0/29) ***Early lucency/subsidence (lumbar only)*** 62% (23/37) vs. 10% (4/41) levels (ALIF mean subsidence: 27% (13-42%) vs. 15% (*P* = NR)) (TLIF mean subsidence: 24% (13-40%) vs. 12% (11.4-13.8%) (*P* = .018))	NR	n/a
Burkus (2011) Integrated analysis, contains studies evaluating on- and off-label uses of rhBMP2 (InFUSE pivotal trial, (including Burkus 2002 and subset of Burkus 2003), Dimar 2009, and another RCT published in abstract only) **rhBMP2 (n = 1093) vs. ICBG (n = 360)**	NR	*rhBMP2 vs. ICBG* **BMP-2 antibody incidence** (3/3 studies)**:** 3.0% (33/1079) (range, 0.8%, 6.4% per study) vs. 1.8% (6/360) (range, 0, 2.3% per study) (*P* = .297) (f/u not clear) (no effect on fusion; all patients with anti-BMP-2 antibodies had rbdiging bone at 6, 12, and 24 mos.) (similar adverse event rates between patients with vs. without antibody responses to BMP, data NR (*P* ≥ .320)). 12 mos (2/3 studies): 0.4% (3/677) vs. NR **BMP-2 neutralizing antibody incidence** (2/3 studies): 0% (0/816) vs. 0% (0/224) **Bovine collagen antibody incidence** (3/3 studies): 16.5% (180/1093) (range, 12.7%, 18.8% per study) vs. 18.2% (66/360) (range, 12.9, 21.2% per study) (*P* = .538) (f/u not clear) (no effect on fusion; data, P-value NR) (similar adverse event rates between patients with	NR	NR

		vs. without antibody responses to BMP, data NR ($P > .25$)). No antibodies against human collagen were detected in any patient, but it was not clear how many of the studies/patients were evaluated. **Miscarriage:** 0.365% (4/1093) (4 events, one pt went on to have live birth) vs NR - none of the 14 patients who became pregnant had a postive BMP-2 antibody response		
Vaccaro 2004/2005/2008 **RCT** **Lumbar Spine** **rhBMP7 (n=24) vs. ICBG (n=12)**	*rhBMP7 vs. ICBG* NR	12 mos.: % patients 24 mos.: total # events (# events ≥ 24 mos.) 48 mos.: total # events (# events ≥ 24 mos.) *rhBMP7 vs. ICBG* **1 yr follow-up:** **All adverse events**: 79% (19/24) vs. 83% (10/12) ($P = 1.0$) **Body as a whole:** 12 mos.: 21% (5/24) vs. 33% (4/12) 24 mos.: NR 48 mos.: NR **Blood and lymphatic system:** 12 mos.: NR 24 mos.: NR 48 mos.: 1 (1) vs. 2 (0) **Cardiac:** 12mos.: NR 24 mos.: 2 (0) vs. 0 (0) 48 mos.: 2 (0) vs. 0 (0) **Cardiovascular:** 12 mos.: 17% (4/24) vs. 17% (2/12)	*rhBMP7 vs. ICBG* None at 1 year follow up. None at 2 years follow up. **4 year follow up:** **revision decompression:** 1 vs. 0 **lumbar decompression and fusion (non-revision):** 1 vs. 0	**Donor site pain (ICBG group only assessed)** 6 weeks (mean): none = 42% (5/12) mild = 33% (4/12) moderate = 25% (3/12) severe = 0% (0/12) 3 mos. (mean): none = 27% (3/11) mild = 55% (6/11) moderate = 18% (2/11) severe = 0% (0/11) 6 mos. (mean): none = 17% (2/12) mild = 50% (6/12) moderate = 17% (2/12) severe = 17% (2/11) 9 mos. (mean): none = 22% (2/9) mild = 33% (3/9) moderate = 44% (4/9) severe = 0% (0/9) 12 mos. (mean): none = 40% (4/10) mild = 40% (4/10) moderate = 10% (1/10) severe = 10% (1/10) 24 mos. (mean): none = 33%

		24 mos.: NR 48 mos.: NR **Digestive/gastrointestinal system:** 12 mos.: 8% (2/24) vs. 17% (2/12) 24 mos.: 2 (0) vs. 3 (1) 48 mos.: 2 (0) vs. 3 (1) **Ear and labyrinth:** 12 mos.: NR 24 mos.: NR 48 mos.: 1 (1) vs. 0 (0) **Eye:** 12 mos.: NR 24 mos.: NR 48 mos.: 2 (2) vs. 0 (0) **General and administration site conditions:** 12 mos.: NR 24 mos.: NR 48 mos.: 2 (1) vs. 4 (1) **Hemic and lymphatic:** 12 mos.: 4% (1/24) vs. 17% (2/12) 24 mos.: NR 48 mos.: NR **Hepatobiliary disorders:** 12 mos.: NR 24 mos.: NR 48 mos.: 1 (1) vs. 0 (0) **Infections and infestations:** 12 mos.: NR 24 mos.: 5 (0) vs. 1 (0) 48 mos.: 6 (1) vs. 1 (0) **Injury, poisoning, and**		mild = 22% moderate = 44% severe = 0%

		procedural complications: 12 mos.: NR 24 mos.: NR 48 mos.: 16 (3) vs. 14 (2) **Investigations:** 12 mos.: NR 24 mos.: NR 48 mos.: 1 (0) vs. 1 (0) **Musculoskeletal and connective tissue:** 12 mos.: 33% (8/24) vs. 25% (3/12) 24 mos.: NR 48 mos.: 40 (11) vs. 21 (12) **Neoplasms, benign, malignant, and unspecified:** 12 mos.: NR 24 mos.: NR 48 mos.: 3 (3) vs. 1 (0) **Nervous system‡:** 12 mos.: 13% (3/24) vs. 8% (1/12) 24 mos.: 2 (0) vs. 3 (1) 48 mos.: 2 (0) vs. 3 (1) **Neurological disorders:** 12 mos.: NR 24 mos.: NR 48 mos.: 1 (0) vs. 0 (0) **Renal and urinary:** 12 mos.: NR 24 mos.: 4 (3) vs. 2 (0) 48 mos.: 1 (0) vs. 0 (0) **Respiratory, thoracid, and mediastinal:** 12 mos.: 0% (0/24) vs. 0% (0/12)		

		24 mos.: 1 (0) vs. 0 (0) 48 mos.: 1 (0) vs. 0 (0) **Skin and appendages:** 12 mos.: 25% (6/24) vs. 0% (0/12) 24 mos.: NR 48 mos.: NR **Skin and subcutaneous tissue:** 12 mos.: NR 24 mos.: NR 48 mos.: 2 (0) vs. 0 (0) **Surgical and medical:** 12 mos.: NR 24 mos.: NR 48 mos.: 3 (2) vs. 0 (0) **Vascular:** 12 mos.: NR 24 mos.: NR 48 mos.: 2 (1) vs. 1 (0) **Ectopic bone formation:** 12 mos.: 0% (0/24) vs. 0% (0/12) 24 mos.: 0 vs. 0 48 mos.: 0 vs. 0 **Recurrent spinal stenosis:** 12 mos.: 0% (0/24) vs. 0% (0/12) 24 mos.: 0 vs. 0 48 mos.: NR **Systemic toxicity:** 12 mos.: 0% (0/24) vs. 0% (0/12) 24 mos.: 0 vs. 0 48 mos.: NR		

		Presence of straight leg tension sign causing pain: Preop: 29% (7/24) vs. 8% (1/12) 6 weeks (mean): 13% (3/24) vs. 0% (0/12) 3 mos. (mean): 13% (3/24) vs. 8% (1/12) 6 mos. (mean): 13% (3/24) vs. 8% (1/12) 9 mos. (mean): 13% (3/24) vs. 0% (0/6) 12 mos. (mean): 5% (1/22) vs. 9% (1/11) 24 mos. (mean): 0% (0/19) vs. 18% (2/11) "There were no complications or adverse events directly related to the OP-1 Putty (rhBMP7), with the possible exception of pseudarthrosis."		
Vaccaro, Lawrence (2008)/ Hwang 2010 **RCT** **Lumbar spine** **rhBMP7 (n = 208 treated) vs. ICBG (n = 87 treated)**	NR	*rhBMP7 vs. ICBG* **Absence of treatment-related Serious Adverse Events (SAEs):** 24 mos.: 85.6% (166/194) vs. 84.7% (61/72) (*P* = .863) 36+ mos.: 79.5% (132/166) vs. 73.5% (50/68) (*P* = .387) **Elevated anti-rhBMP7 antibodies (any time point; 6 weeks,3, 6, 12, and 24 mos.):** 93.7% vs. 20.9% **Positive for anti-rhBMP7 neutralizing antibodies:** 25.6% vs. 1.2% (peak for neutralizing antibodies between 6 weeks and3 mos.; at 24 & 36+ mos. no patients positives for	*rhBMP7 vs. ICBG* **Revision:** 36 mos.: 8.2% (21/257) vs. 13% (11/87) 36-48+ mos.: 2.1% (3/144) vs. 5.2% (3/58%) (*P* = .242)	**Donor site pain (VAS):** *rhBMP7:* NR *ICBG:* *36+* mos.: 35% patients reported mild/moderate pain

		neutralizing antibodies) "No significant associations were observed between neutralizing activity status, clinical success, and safety parameters."		
Johnsson (2002) **RCT** **Lumbar spine** **rhBMP7 (n = 10) vs. ICBG (n = 10)**	"No intraoperative complications occurred."	"No early, late, local, or systemic adverse effects of the OP-1 (rhBMP7) Implant were noted."	*rhBMP7 vs. ICBG* **Decompression:** 10% (1/10) vs. 10% (1/10) **Instrumented fusion:** 20% (2/10) vs. 0% (0/10)	*rhBMP7 vs. ICBG* **Iliac crest pain (1 year):** 0% (0/10) vs. 10 (1/10)
Kanayama (2006) **RCT** **Lumbar spine** Posterolateral lumbar fusion with pedicle screw instrumentation **rhBMP7 (n = 9) vs. local HT-TCP/ autograft (n = 10)** (HT-TCP: hydroxyapatite/tricalcium phosphate biphasic ceramic granules; ceramic bone substitute) (rhBMP7 group had local autograft taken but was not used)	NR	NR	NR	NR
Delawi et al. (2010) RCT **Lumbar spine** Treatment groups: *rhBMP7*: n = 18 *autograft*: n = 16	*rhBMP7 vs. autograft* Dural tear: 5.6% (1/18) vs. 6.3% (1/16) Surgical infection: 5.6% (1/18) vs. 6.3% (1/16) Hematoma: 11.1% (2/18) vs. 0% (0/16) Neural injury: 5.6% (1/18) vs. 6.3% (1/16) Herniation: 5.6% (1/18) vs. 0% (0/16) *P* = ns for all	*rhBMP7 vs. autograft* Cardiovascular: 5.6% (1/18) vs. 6.3% (1/16) Respiratory: 5.6% (1/18) vs. 0% (0/16) Malignancy: 5.6% (1/18) vs. 0% (0/16) Instrumentation failure: 0% (0/18) vs. 6.3% (1/16) Excessive leg pain: 5.6% (1/18) vs.12.5% (2/16) *Total complications (surgical and adverse events):* 55.6% (10/18) vs. 43.8%	NR	VAS (1-10; mean ± SD) *6 weeks:* 3.0 ± 2.8 *3 months:* 1.7 ± 1.7 *6 months*: 3.8 ± 3.5 *12 months*: 2.7 ± 2.8 At 12 months, 64% of patients classified their pain as "Mild" "No complications directly related to the bone graft harvesting procedure occurred"

	comparisons	(7/16) *P* = ns for all comparisons		
FDA SSPB for OP-1 HDE H020008 2004	NR	*rhBMP7 (OP-1) vs. autograft* **Abnormal lab values**: 3% (6/228) vs. 8% (8/98) **Blood and lymphatic system disorders:** 4% (8/228) vs. 14% (14/98) **Cardiac disorders:** 4% (9/228) vs. 1% (1/98) **Gastrointestinal disorders:** 13% (30/228) vs. 10% (10/98) **General disorders and admnistration site condition:** 16% (36/228) vs. 18% (18/98) **Infections and infestations:** 8% (18/228) vs. 8% (8/98) **Injury, poisoning and procedural complications:** 19% (44/228) vs. 24% (23/98) **Metabolism and nutrition disorders:** 3% (6/228) vs. 1% (1/98) **Musculoskeletal and connective tissue disorders - other:** 22% (50/228) vs. 24% (23/98) **Musculoskeletal and connective tissue disorders - joint inflammation:** 11% (24/228) vs. 6% (6/98) **Musculoskeletal and connective tissue disorders - pseudarthrosis:** 5% (12/228) vs. 3% (3/98)	NR	NR

		Nervous system disorders - other: 11% (26/228) vs. 10% (10/98) **Nervous system disorders - TIA:** 2% (4/228) vs. 0% (0/98) **Psychiatric system disorders:** 4% (10/228) vs. 3% (3/98) **Renal and urinary disorders:** 6% (13/228) vs. 9% (9/98) **Respiratory, thoracic and mediastinal disorders:** 7% (15/228) vs. 4% (4/98) **Skin and subcutaneous tissue disorders - other:** 4% (8/228) vs. 1% (1/98) **Skin and subcutaneous tissue disorders - wound infection:** 7% (15/228) vs. 2% (2/98) **Vascular disorders:** 8% (17/228) vs. 10% (10/98) **Cancer (worldwide reporting):** 7 cases vs. NR -6 of 7 cases non-osseous cancers -7th case, recurring chondrosarcoma in patient with a history of chondrosarcoma -incidence of cancer in *rhBMP7* patients is in the range of cancer occurrence in general populations **Antibodies detected:** 96% (23/24) vs. 0% **Neutralizing antibodies detected:** 29% (7/24) vs. 0% -6 patients had neutralizing antibodies detected at 6 weeks postop but not at 6 mos. postop		

		-1 patient had neutralizing antibodies detected only at 6 mos. postop **ear & labrinth disorders; eye disorders; immune system disorders; neoplasms (benign, malignant, or unspecified); reproductive system and breast disorders; social circumstances; surgical and medial procedures** Seen in < 1% of investigational population		
Cahill et al. (2009) Retrospective cohort (database) study **Lumbar spine** (subset of total population) Treatment groups: *rhBMP (any)*: n = 13,972 *Non-BMP*: n = 22,835	*BMP vs. No BMP* **Any complication:** 6.97% (974/13,972) vs. 7.18% (1639/22,835) Unadjusted OR: 0.96 (95% CI, 0.89, 1.05) Adjusted§ OR: 1.03 (95% CI, 0.95, 1.12) **Dysphagia or hoarsness:** 0.25% (36/13,972) vs. 0.21% (49/22,835) Unadjusted OR: 1.20 (95% CI, 0.78, 1.84) Adjusted§ OR: not applicable **Wound complication:** 2.01% (281/13,972) vs. 2.15% (507/22,835) Unadjusted OR: 0.90 (95% CI, 0.78, 1.04) Adjusted§ OR: 0.93 (95% CI, 0.80, 1.08) **"Other complications":** 4.98% (696/13,972) vs. 5.12% (1170/22,835)	NR	NR	NR

	Unadjusted OR: 0.97 (95% CI, 0.88, 1.06) Adjusted§ OR: 1.05 (95% CI, 0.95, 1.15)			
Cahill et al. (2011) Retrospective case-control (database) study **Lumbar spine** Treatment groups: *rhBMP (any)*: n = 2372 *Non-BMP*: n = 2372	NR	*BMP vs. No BMP* **Readmission** (within 30 days): 3.9% *vs.* 5.0% (*P* = .08) Unadjusted OR: 0.77 (95% CI, 0.58, 1.02) Multivariate adjusted OR: 0.72 (95% CI, 0.54, 0.95)	*BMP vs. No BMP* **Repeat fusion:** All rates are cumulative. ***1 year:*** 2.3% vs. 3.4% (*P* = .03) Unadjusted OR: 0.65 (95% CI, 0.47, 0.90) Multivariate adjusted (for other significant predictors) OR: 0.66 (95% CI, 0.47, 0.94) ***"Long-term":*** BMP associated with decrease (*P* = .01) (2 yrs: 5.2% vs. 6.6%; 3 yrs: 6.8% vs. 9.2%) Unadjusted HR: 0.75 (95% CI, 0.59, 0.95) Multivariate adjusted (for other significant predictors) HR: 0.74 (95% CI, 0.58, 0.93)	NR
Deyo et al. (2011) Retrospective cohort (database) study **Lumbar spine** Treatment groups: *rhBMP (any)*: n = 1703 *Non-BMP*: n = 15,119	NR	*BMP vs. No BMP* **Readmission** (within 30 days): 12.0% (205/1703) vs. 12.3% (1855/15,119) (*P* = .782**) **Cardiac, pulmonary, or stroke complications:** 5.1% (87/1703) vs. 5.7% (868/15,119) (*P* = .285**)	*BMP vs. No BMP* **Reoperation** (within 6 mos): 1.2% (21/1703) vs. 1.2% (186/15,119) (*P* = .992**) **Reoperation** (within 1 yr): 2.7% (46/1703) vs. 2.9% (443/15,119) (*P* = .594**)	NR

		Wound complications: 2.4% (40/1703) vs. 2.2% (332/15,119) (*P* = .684**) **Death** (within 30 days)**:** 0.9% (15/1703) vs. 0.8% (118/15,119) (*P* = .656**) **Nursing home discharge:** 15.9% (271/1703) vs. 19.0% (2869/15,119) (*P* < .001**)	**Reoperation** (within 2 yrs): 6.3% (107/1703) vs. 6.0% (912/15,119) (*P* = .681**) **Reoperation** (within 3 yrs): 9.2% (157/1703) vs. 8.5% (1287/15,119) (*P* = .324**) **Reoperation** (within 4 yrs): 10.8% (183/1703) vs. 10.5% (1588/15,119) (*P* = .757**)	
Mines et al. (2011) Retrospective cohort (database) study **Lumbar spine** **rhBMP-2:** n = 15,460 **No BMP:** 78,194	NR	*rhBMP-2 vs. No BMP:* **Pancreatic cancer:** 0.052% (8/15,460) vs. 0.106% (83/78,194) (OR: 0.49 (95% CI,0.24, 1.02 (univariate analysis); BMP use was not associated with pancreatic cancer in either unadjusted (HR: 0.68 (95% CI, 0.33, 1.42) or multivariate (HR: 0.70 (95% CI, 0.34, 1.45)) Cox regression analysis. **Death:** 3.1% (479/15,460) vs. 5.1% (2988/78,194) (*P* = NR)	NR	NR
Baskin et al. (2003) RCT **Cervical spine – DDD** **rhBMP-2** (1.5 mg/mL; 0.4 mL reconstituted) with CORNERSTONE-SR allograft ring and ATLANTIS cervical plate, n = 18 **ICBG** with CORNERSTONE-SR	*rhBMP-2 vs. ICBG* "no unanticipated device-related adverse events in either treatment group"	**Positive antibody response to rhBMP-2**: no patient in either group **Formation of ectopic bone anterior to the spine at an adjacent level:** rhBMP-2: n = 2 (11.1%) ICBG: n = 1 (6.7%) "The number of patients in this study is too small to assess whether BMP may increase the rate of ectopic bone formation in	**rhBMP-2**, n = 1 (5.6%); adjacent segment to the original 2-level fusion, unrelated to original procedure **ICBG**, n = 0	6 weeks postop, ICBG patients had significant levels of pain at graft site (*P* < .007) and complained about appearance of graft site 6 months postop no statistical differences between groups in terms of graft-site pain or appearance At 24 month follow-

allograft ring and ATLANTIS cervical plate, n = 15		this clinical application. This issue should be investigated further."		up, some ICBG patients continued to experience residual pain and rate appearance of site as only fair
Butterman (2008) Prospective cohort **Cervical spine –** DDD, HNP, stenosis **rhBMP-2** (0.9 mg/level) with allograft, n = 30 **ICBG**, n = 36	*rhBMP-2 vs. ICBG* **Neck problems** ***Swelling (1º complaint new on-set dysphagia)*** 50.0% (15/30) vs. 13.9% (5/36) - In rhBMP group, symptoms occurred at mean 4 ± 3 days postop and lasted 21 ± 16 days - occurred most often in 2-level fusion (62.5%, 10/16); 1-level fusion (50.0%, 2/4); 3-level fusion (30.0%, 3/10) ***Re-admit*** 10.0% (3/30)†† vs. 0% (0/36) ***MD evaluation*** 23.3% (7/30) vs. 8.3% (3/36) ***Phone call – RN*** 33.3% (10/30) vs. 11.1% (4/36)	*rhBMP-2 vs. ICBG* **Pseudarthrosis** 3.3% (1/30) vs. 5.6% (2/36) *P* = ns **Delayed union** 0% (0/30) vs. 2.7% (1/36) **Adjacent segment disc herniation above fusion, at 2 years postop** 3.3% (1/30) vs. 5.6% (2/36) "**neurological deficits** (weakness, altered senstation) uniformly resolved in both groups"	**rhBMP-2:** 3.3% (1/30) – 1 ACDF extension with decompression for adjacent segment disc herniation above fusion **ICBG:** 8.3% (3/36) – 1 irrigation and debridement of graft site infection; 1 ORIF of ASIS fracture; 1 pseudarthrosis repair with single level posterior instrument fusion	**1 year, VAS pain at graft site (0–10)**: 0.2 ± 0.7 **Infection:** 2.7% (1/36) **ASIS fracture:** 2.7% (1/36)
Crawford et al. (2009) Retrospective cohort **Cervical spine** – stenosis, ACDF nonunion, spondylosis **rhBMP-2** (mean 3.6 mg per level), n = 41 **ICBG**, n = 36	*rhBMP-2 vs. ICBG* **Medical complications** 0% (0/41) vs. 8.3% (3/36) *P* = ns ***Postop tachycardia:*** 0% vs. 2.8% (1/36) ***Transfusion for postop anemia:*** 0% vs. 2.8 (1/36) ***Nausea, vomiting, and headaches :*** 0% vs. 2/8% (1/36)	*rhBMP-2 vs. ICBG* **Prolonged wound drainage** 4.9% (2/41) vs. 2.8% (1/36) *P* = ns **Deep infection** 9.8% (4/41) vs. 0% (0/36) *P* = ns - rhBMP dose for those with infection vs. without infection: 2.9 mg/level vs. 3.7 mg/level; *P* = ns	**rhBMP-2:** 9.8% (4/41) – all had irrigation and debridement with IV antibiotics for deep infections **ICBG:** 2.8% (1/36) - irrigation and debridement with IV antibiotics for deep infection of iliac crest site	**Deep infection of iliac crest site**, 2.7% (1/36)
Smucker et al (2006) Retrospective cohort (chart review with concurrent control) **Cervical spine –**	*rhBMP-2 vs. ICBG* **Swelling complications** ***Total:*** 27.5% (19/69) vs. 3.6% (6/165); *P* < .0001	NR	See surgical and perioperative complications	NR

rhBMP-2 (1.5 mg/mL): n = 69 **ACDF** with allograft or autograft: n = 165	- in rhBMP-2 group, swelling occurred at a mean 4.2 (range, 2–7) days postop in 11/19 patients in whom onset could be determined ***Delay in discharge due to:*** **Visible neck swelling:** 2.9% (2/69) vs. 0% (0/165) **Severe dysphagia:** 7.2% (5/69) vs. 1.2% (2/165) **Reintubation:** 2.9% (2/69) vs. 0% (0/165) **PEG placement:** 1.4% (1/69) vs. 1.2% (2/165) **Tracheostomy** 1.4% (1/69) vs. 0.6% (1/165) **Delay in extubation** 0% (0/69) vs. 0.6% (1/165) ***Incision and drainage of swollen surgical site*** 4.3% (3/69) vs. 0% (0/165) ***Readmit for medical management of swelling*** 2.9% (2/69) vs. 0% (1/165) ***Premature return to clinic or ER visit*** 4.3% (3/69) vs. 0.6% (1/165) ***Outpatient ENT consult*** 2.9% (2/69) vs. 0% (0/165) Multivariate regression showed that rhBMP-2 usage remained significantly associated with cervical swelling complications (*P* < .0001); adjusted OR =			

	10.1 (95% CI, 3.8, 26.6)			
Vaidya, Carp et al. (2007) Retrospective cohort **Cervical spine** – DDD, stenosis **rhBMP-2** with PEEK cages, n = 22 **allograft** and demineralized bone matrix with plate, n = 24	*rhBMP-2 vs. allograft* **Hospital stay** 2.9 (1–9) vs. 2.3 (1–6) days - In rhBMP-2 group, stay prolonged by 3 patients with "severe" dysphagia	*rhBMP-2 vs. allograft* **Nonunion** 0% (0/22) vs. 4.2% (1/24) **Suspected infection** 4.5% (1/22) vs. 0% (0/24) **Continued neck pain in the upper cervical spine** 4.5% (1/22) vs. 0% (0/24) **Dysphagia** *Overall:* 90.9% (20/22) vs. 75.0% (18/24) *At 2 weeks*: 85.0% (17/20) vs. 38.9% (7/18) *P* = .01 *At 6 weeks*: 65.0% (13/20) vs. 22.2% (4/18) *P* = .02 *Single level fusion, at 2 weeks:* 71% (16/22) vs. 13% (3/24) *P* = .07 *2- and 3-level fusion, 6 weeks:* 92% (20/22) vs. 40% (10/24) *P* = .02 *At 2 years*, 21% of patients still complained of dysphagia (20% rhBMP-2; 22% allograft) **Hoarseness of voice** *Postop:* 60% (13/22) vs. 62% (15/24); *P* = ns *Last follow-up:* 9.1% (2/22) vs. 12.5% (3/24); *P* = ns 1-level, 2- or 3-level cases were all similar between groups.	**rhBMP-2:** 9.1% (2/22) - 1 wound exploration for suspected infection early postop; 1 operation at a lower level **Allograft:** 4.2% (1/24) Revision surgery for nonunion at 12 months	NR
Vaidya, Weir et al. (2007) **Prospective cohort** **Cervical (+ lumbar, NR here) spine**	NR	*rhBMP2/allograft vs. DBM/allograft* ***Nonunion (cervical only)*** 0% (0/11) vs 8% (1/12) (required reoperation (plate removal & posterior fusion) ***Early lucency/subsidence (cervical only)*** 62% (6/18) vs. 0% (0/22)	NR	n/a

		levels (mean subsidence: 53% (40-58%) vs. <10%)		
Xu (2011) Retrospective cohort **Cervical spine** **rhBMP2 (n = 48) vs. non-BMP2 (n = 156)**	*rhBMP2 vs. non-BMP2* **Incidental durotomy:** 0% (0/48) vs. 2.6% (4/156) (*P* = .26) **CSF leakage:** 0% (0/48) vs. 1.3% (2/156) (*P* = .43) Follow up interval: 24.2 ± 10.1 mos.(1-39.6 mos.)	*rhBMP2 vs. non-BMP2* **Incidental durotomy:** 0% (0/48) vs. 2.6% (4/156) (*P* = .26) **CSF leakage:** 0% (0/48) vs. 1.3% (2/156) (*P* = .43) **Deep vein thrombosis:** 0% (0/48) vs. 1.9% (3/156) (*P* = .33) **Pulmonary embolism:** 0% (0/48) vs. 1.3% (2/156) (*P* = .43) **Hyperostosis:** 0% (0/48) vs. 0% (0/156) (*P* = 1) **Infection:** 10.9% (5/48) vs. 10.9% (17/156) (*P* = .93) **Pneumonia:** 2.2% (1/48) vs. 2.0% (4/156) (*P* = .85) **Dysphagia:** 6.3% (3/48) vs. 3.8% (6/156) (*P* = .48) **Hematoma:** 2.2% (1/48) vs. 1.9% (3/156) (*P* = .94) **C5 palsy:** 6.5% (3/48) vs. 4.5% (7/156) (*P* = .62) **Wound dehiscence:** 2.2% (1/48) vs. 5.1% (8/156) (*P* = .37) **Instrumentation failure:** 0% (0/48) vs. 7.1% (11/156) (*P* = .06)	*rhBMP2 vs. non-BMP2* **Reoperation:** 15.2% (7/48) vs. 20.5% (32/156) (*P* = .36) Follow up interval: 24.2 ± 10.1 mos.(1-39.6 mos.)	NR

		Discharge to rehabilitation: 28.3% (13/48) vs. 35.4% (55/156) (*P* = .29) Follow up interval: 24.2 ± 10.1 mos.(1-39.6 mos.)		
Cahill et al. (2009) Retrospective cohort (database) study **Cervical spine** (subset of total population) Treatment groups: ***Anterior cervical*** *rhBMP (any)*: n = 2299 *Non-BMP*: n = 24,768 ***Posterior cervical*** *rhBMP (any)*: n = 478 *Non-BMP*: n = 2391	*BMP vs. No BMP* ***Anterior cervical*** **Any complication:** 7.09% (163/2299) vs. 4.68% (1158/24,768) Unadjusted OR: 1.55 (95% CI, 1.31, 1.84) Adjusted§ OR: 1.43(95% CI, 1.20, 1.70) **Dysphagia or hoarsness:** 4.35% (100/2299) vs. 2.45% (608/24,768) Unadjusted OR: 1.80 (95% CI, 1.45, 2.24) Adjusted§ OR: 1.67 (95% CI, 1.30, 2.05) **Wound complication:** 1.22% (28/2299) vs. 0.65% (160/24,768) Unadjusted OR: 1.89 (95% CI, 1.26, 2.83) Adjusted§ OR: 1.67 (95% CI, 1.10, 2.53) **"Other complications":** 2.39% (55/2299) vs. 1.94% (480/24,768) Unadjusted OR: 1.25 (95% CI, 0.93, 1.64) Adjusted§ OR: 1.16 (95% CI, 0.87, 1.56) ***Posterior cervical*** **Any complication:** 10.04% (48/478) vs. 9.95% (238/2391)	NR	NR	NR

	Unadjusted OR: 1.01 (95% CI, 0.72, 1.40) Adjusted§ OR: 1.03 (95% CI, 0.73, 1.44) **Dysphagia or hoarsness:** 2.09% (10/478) vs. 1.63% (39/2391) Unadjusted OR: 0.59 (95% CI, 0.24, 1.41) **Wound complication:** 2.93% (14/478) vs. 2.51% (60/2391) Unadjusted OR: 1.17 (95% CI, 0.64, 2.11) Adjusted§ OR: 1.11(95% CI, 0.60, 2.05) **"Other complications":** 5.86% (28/478) vs. 6.48% (155/2391) Unadjusted OR: 0.89 (95% CI, 0.59, 1.35) Adjusted§ OR: 0.94 (95% CI, 0.61, 1.44)			
Yaremchuk (2010) **Retrospective cohort study** **Cervical spine** **BMP (n = 260) vs. non-BMP (n = 515)**	See Adverse Events	*rhBMP2 vs. non-BP2* **Death:** 4.2% (11/260) vs. 1.7% (9/515) (within 90 d post-surgery) (*P* = .047) **Percutaneous endoscopic gastrostomy (PEG):** 42.3% (6/260) vs. 0.8% (4/515) (within 30 d post-surgery) (*P* = .089) **Tracheotomies:** 3.1% (8/260) vs. 0.6% (3/515) (within 30 d post-surgery) (*P* = .024) **Unplanned intubations**	NR	NR

		after surgery: 6.2% (16/260) vs. 1.6% (8/515) (within 30 d post-surgery) (*P* = .008) **Readmissions:** 8.8% (23/260) vs. 5.0% (26/515) (within 30 d post-surgery) (*P* = .040) **Dysphagia:** 6.9% (18/260) vs. 3.3% (17/515) (within 30 d post-surgery) (*P* = .001) **Dyspnea:** 20.4% (53/260) vs. 8.0% (41/515) (within 30 d post-surgery) (*P* = .001) **Hoarseness:** 2.3% (6/260) vs. 1.2% (6/515) (within 30 d post-surgery) (*P* = .427) **Respiratory failure:** 13.1% (34/260) vs. 4.7% (24/515) (within 30 d post-surgery) (*P* = .001)		

ACDF: anterior cervical discectomy and fusion; ASF: anterior spinal fusion; ASIS: anterior superior iliac spine; CBC: complete blood count; CSF: cerebrospinal fluid; DDD: degenerative disc disease; DVT: deep vein thrombosis; ENT: ear, nose and throat; ER: emergency room; GI: gasterointestinal; HNP: herniated nucleus pulpous; HR: hazards ratio; ICBG: iliac crest bone graft; IVC: inferior vena cava; n/a: not applicable; MD: medical doctor; NR: not reported; OR: odds ratio; ORIF: open-reduction and internal fixation; rhBMP-2: recombinant human bone morphogenetic protein-2; RE: retrograde ejaculation; RN: registered nurse ;tx: treatment; UTI: urinary tract infection.
*20-point VAS scale derived from the summation of the numeric rating scores for pain intensity and pain duration. Higher scores = greater pain.
†AHRQ reported this as a prospective study.
‡Discrepancy between 12 mos. and 24/48 mos. follow-up reporting.
§Adjusted for age, race, sex, income, elective admission, teaching hospital, revision surgery, diagnosis, medical comorbidities, levels fused, primary payer, and geographic location of hospital.
**Similar p-values following regression analysis to adjust for baseline differences in age, sex, race, comorbidity score, previous hospitalizations without spine surgery, previous spine surgery, previous hospitalizations, simple or complex fusion, and presence of spondylolisthesis or scoliosis.
††The 3 readmissions were due to neck swelling causing dysphagia; admitted to intensive care unit for observation and treated with IV steroids (none required additional surgery, however).

Appendix Table 12. Case series evaluating the safety of BMPs in lumbar spinal fusion.

Investigator (yr, country, ref #) Surgical Site	No. pts Sex (% male) Mean age (BMP dose)	Diagnosis	Surgical intervention	Follow-up: Mean duration Loss to f/u (%)	Reported complications
On-label use: rhBMP-2					
Burkus 2009 (note: 6-yr follow-up data for the BMP patients reported in Burkus 2002[14] and the FDA InFUSE SSED[61].	N = 277 47.9% male* 43.2 years*	Single-level DDD	1-level open or laproscopic ALIF	2 years 80.1% (222/277) 6 years 52.7% (146/277)	Second surgery (any): Cumulative (> 6 years): 10.4% (25/277); with rate adjusted based on the number of patients available at each follow-up interval using a time-to-event analysis. -23 supplemental fixations, 1 cage removal, 1 revision 2 years: 8.1% (18/222) 2-6 years: 4.8% (8/277) Anatomical and/or technical difficulty: 2 years: 4.1% (9/222) (0/9 required second surgical procedure) 2-6 years: 0% (0/277) Malpositioned implant: 2 years: 2.3% (5/222) (2/5 required second surgical procedure) 2-6 years: 0% (0/277) Implant displacement/loosening: 2 years: 1.8% (4/222) (1/4 required second surgical procedure) 2-6 years: 0% (0/277) Subsidence: 2 years: 3.2% (7/222) (4/7 required second surgical procedure) 2-6 years: 0% (0/277)

Off-label use: rhBMP-2					
Anderson 2011	N = 50 52% male 48.2 years (32-84) Dosage: NR	Degenerative spine disease	1- or 2-level ALIF (mean 1.4 levels/pt)	12 months 0% (0/50) loss to f/u†	Intraoperative complications 0% (0/50) Total postoperative complication rate 12% (6/50) Ileus requiring an NG tube for 2 days 2% (1/50) Scrotal edema 2% (1/50) Tachycardia, transient hypotension with trace pericardial effusion (medically managed) 2% (1/50) Urinary retention 2% (1/50) Urinary tract infection 4% (2/50) Wound infection 0% (0/50) Thromboembolic disease 0% (0/50) Symptomatic pseudoarthrosis 0% (0/50) Hardware loosening or failure 0% (0/50) Hardware repositioning 0% (0/50)

Carreon 2008	N = 96 46% male Mean age: NR Dosage: NR	NR (Comparison of patients with 2 spinal surgeries)	1st surgery: 28 cervical 3 thoracic 65 lumbar 1.9 ± 1.2 levels fused 2nd Surgery: 24 cervical 5 thoracic 67 lumbar 2.2 ± 1.7 levels fused	NR 0% (0/96) loss to f/u	Total complications 1st surgery: 44 (in 38 patients) 2nd surgery: 30 (in 27 paatients) Deep wound infections requiring multiple debridements 1st surgery: 2% (2/96) 2nd surgery: 4% (4/96) Wound drainage or hematomas (did not require surgical intervention 1st surgery: 9% (9/96) 2nd surgery: 11% (11/96) Allergic reactions (anaphylactic) 1st surgery: NR 2nd surgery: 0% (0/96)
Garrett 2010	N = 130 96% male 58 years (34-80) Dosage: 8.4 mg/patient (2.1-14.7mg)	NR	Posterolateral lumbar fusion Mean 3.5 levels (1-8)	NR 0% (0/130) loss to f/u	Durotomy 2% (3/130) (2 cases required direct repair) Painful seroma and edema 4.6% (6/130)
Geibel 2009	N = 48 52% male 49.7 ± 9.6 years (males) 50.6 ± 8.6 years (females) Dosage: NR	Degenerative disk disease (25% grade I isthmic spondylolisthesis)	Posterior lumbar interbody fusion (PLIF) Mean 1.2 levels	16.9 (11.2-23.8) months 0% (0/48) loss to f/u	Central canal compromise 0% (0/48) Adjacent level fusion 0% (0/48) Heterotopic bone formation 0% (0/48)

Glassman 2010/2011	N = 1037 38.6% male 58.4 (18-90) years Dosage: 12-24 mg **2011** N = 109 35.7% male	Diagnosis (cases) Stenosis (253) Spondylolisthesis (204) Instability (22) Scoliosis (29) Disc pathology (106) Nonunion (115) Adjacent level degeneration (180) postdiscectomy instability (128)	Posterlateral fusion Mean 1.8 (1-5) levels	3 months 0% (0/1037) loss to f/u **2011** 6.4% (7/109) loss to f/u	Total medical and surgical complications: 18.3% (190/1037 patients) Major complications 7.8% (81/1037) Pneumonia 1.64% (17/1037) Respiratory failure 0.29% (3/1037) Pulmonary embolism 0.10% (1/1037) Myocardial infarction 0.19% (2/1037) Arrhythmia 0.58% (6/1037) Cardiac ischemia 0.10% (1/1037) Acute renal failure 0.19% (2/1037) Urosepsis 0.29% (3/1037) Pulmonary embolism 0.10% (1/1037) Other 1.16% (12/1037) Deep wound infection 2.12% (22/1037) Hematoma (neg. culture) 0.96% (10/1037) Screw malposition 0.58% (6/1037) Epidural hematoma 0.29% (3/1037) Retained drain 0.10% (1/1037) Excessive blood loss 0.29% (3/1037)

					Radiculopathy 0.68% (7/1037) Minor complications 10.2% (110/1037) Psoas hematoma 0.77% (8/1037) Superficial wound infection 1.74% (18/1037) Urinary tract infection 1.83% (19/1037) Ileus 2.60% (27/1037) Mental status change 3.66% (38/1037) Dural tear 5.59% (58/1037)
Helgeson 2011	N = 88 (65 patients excluded due to lack of imaging at required postop times); 23 patients met inclusion criteria 78% male (18/23) 38.2 (23-81) years Dosage: 6 mg	NR	TLIF Mean 1.7 (1-3) levels	1-2 years 74% (65/88) loss to f/u	Osteolysis (incidence in adjacent vertebral bodies) 3-6 mos.: 54% (specific data NR) 1-2 years: 41% (specific data NR)

Knox 2011	N = 71 (10 patients excluded due to lack of imaging, 2 excluded due to incomplete operative documentation, 1 excluded due to postop infection); 58 patients included 72% male (42/58) 36.9 (20-61) years Dosage: 5 mg/level	Degenerative spinal conditions (spondylolis-thesis, discogenic back pain, lumbar radiculopathy)	TLIF with pedicle screw instrumentation Mean 1.3 (1-2) levels	4.3 (2.4-9) months 18% (13/71) loss to f/u	Osteolysis 27% (16/58) patients 26% (20/77) levels 21% (8/39) patients with single-level fusion 50% (8/19) patients with two-level fusion Graft Subsidence 10% (6/58) patients 8% (6/77) levels (evidence of subsidence was not evident on the initial postoperative CT) all incidences of graft subsidence occurred with severe osteolysis Migration of intervertebral cage 9% (5/58) patients

Luhmann 2005	N = 70 (95 procedures)
	20% male
	55 years
	
	<u>Mean doses/level:</u>
	ALIF 10.8mg
	Posterior 13.7mg
	Compassionate use 28.6mg
		Spinal deformity	ALIF 48% procedures (46/95)
			
			Posterior 43% procedures (41/95)
			
			Compassionate use 8% procedures (8/95)
				17.9 (12–60) months
				
				% f/u NR
					• Deep wound hematoma:
					  • 1% (1/70) (no long-term clinical sequelae)
					
					• Wound infection or dehiscence: 3% (2/70)
					  • Deep wound infection (n = 1)
					  • Superficial wound infection (n = 1)
					
					• Toxicity (local or systemic)
					  • 0% (0/70)

Mannion 2011	N = 30 47% male (14/30) 53 (22-78) years Dosage: 1.4 mg/level	Central canal stenosis, foraminal stenosis/colla-pse, discogenic back pain and disc prolapse	TLIF 89% (32/36) levels PLIF 11% (4/36) Mean 1.2 (1-2) levels	7.1 months % f/u NR	Heterotopic ossification 7% (2/30) patients Inflammatory cyst in the neural foramen 7% (2/30) patients Cage subsidence 3% (1/30) patients Osteolysis 3% (1/30) patients
McClellan 2006	N = 26 46% female 46 Years BMP doses variable and not controlled.	1-2 level DDD. Radiculopat hy present in some cases.	TLIF with rhBMP-2/ACS. A variety of allografts and interbody fusion cages were used.	CT scans at 3-7 months (mean 4.4). % f/u NR	• Vertebral resorption (clinical relevance not investigated) • 69% (22/32) lumbar levels. • This was characterized as: • Mild: 50% (11/22) • Moderate: 18% (4/22) • Severe: 32% (7/22) • Graft subsidence or loss of endplate integrity (clinical relevance not investigated) • 16% (5/32) • 5/5 had severe vertebral resorption
Meisel 2008	N = 17 47% male 67 Years Doses: 12 mg/level. 6 mg/level for one patient who had a 2-level fusion	Lumbar DDD with stenosis and invertebral instability.	1-2 level PLIF with rhBMP-2/ACS-filled PEEK cage.	3, 6, 12, 24, and 36 months % f/u NR	• Transient bone resorption (no effect on clinical success) • 100% (17/17) patients • Detencted at 3 months with ossification observed at 6 months • Patients asymptomatic • Intracanal bone formation • 6% (1/17) patients • Patient asymptomatic

Mindea 2009	N = 35 42% male 51 Years[i] Doses: 4.2 mg/level	Grade I or II Spondylolis-thesis, mechanical back pain, or recurrent disc herniation.	Minimally invasive single-level thoracic with rhBMP-2/ACS, as well as autograft and pedicle screws.	NR % f/u NR	• Radiculitis • 11% (4/35) • New onset postoperatively. Patients had no structural evidence of radiculitis (CT).

Owens 2011	N = 204 44.6% male 49.3 (22-79) years Doses: 1.4-6 mg (96% of patients had 4 mg)	Spondylolisth-esis, instability, stenosis, scoliosis, disc pathology, nonunion, adjacent level degeneration, post-discectomy instability	TLIF with rhBMP-2 Mean 1.2 (1-2) levels	29.8 ± 9.0 months % f/u NR	Total complications 21.6% (47/204) patients Major complications 6.4% (13/204) patients Pneumonia 0.5% (1/204) patients Vascular Injury 0.5% (1/204) patients Neurologic 3.4% (7/204) patients Wound Infection 1.5% (3/204) patients Wound hematoma/seroma 0.5% (1/204) patients Seroma in the foramen 2.0% (4/204) patients Minor complications 16.7% (34/204) patients Radiculopathy-CT 2.9% (6/204) patients Superficial wound dehiscence 1% (2/204) patients Ileus 2.9% (6/204) patients Urinary tract infection 1% (2/204) patients Other 8.8% (18/204) patients

Sethi 2011 (lumbar, cervical)	N = 95
	55% male
	51 (18-79) years
	
	Dosage:
	2 mg/level in lumbar spine
	1 mg/level in cervical spine
			ALIF (23 patients)
			TLIF (36 patients)
			PLIF (2 patients)
			Anterior cervical decompression and fusion (34 patients)
			
			Mean 1.4 levels
			
			Polyetherether-ketone cage used in 59 patients (82 levels)
				2, 6 weeks, 3, 6, 12, and 24 months	• End plate resorption **(lumbar)**
					  • 82% (71/87) levels in lumbar spine showed some resorption
					  • 18% (16/87) levels had no resorption at all
					  • Largest transition was at 6 to 9 months post-op
					• Subsidence/narrowing of disk space **(lumbar + cervical)**
					  • 50% of patients (47/95)
					  • Average subsidence for group was 16.5% at 12 months
					• Heterotopic bone formation **(lumbar)**
					  • Stated as "commonly seen in TLIF patients" but data is NR
					  • Symptoms tended to appear 6-8 weeks after surgery
					  • NR for both ALIF and PLIF procedures
					• Cage migration **(lumbar + cervical)**
					  • Lumbar: 10-11/61 patients, with 10 of which underwent TLIF with a PEEK cage.
					  • TLIF with PEEK cage: 38% (10/36)
					  • Cervical: 0-1/34 patients
					  • Unclear if the one additional case occurred in lumbar or cervical.

Stambough 2011	N = 36 22% male mean age 66.3 (34-87) years <u>Doses:</u> 12 mg	Lumbar acquired spinal stenosis, degenerative disc disease	Mean 1.44 (1-2) levels Posterolateral fusion with rhBMP-2 and allograft	28.6 (24-34) months 0% (0/36) loss to f/u	Dural tear 3% (1/36) Infection 0% (0/36)

Vaidya 2008 (cervical + lumbar)	N = 59 (82 levels)
	
	Lumbar fusions:
	N = 36 (50/82 levels)
	
	Cervical Fusions:
	N = 23 (32/82 levels)
	
	% male NR
	52 years
	
	<u>Doses</u>:
	2 mg/level
		Spondylolis-thesis, adult scoliosis, revision surgery, discogenic pain	Single- or multiple- level lumbar (ALIF, PLIF, TLIF) spinal fusions with rhBMP-2/ACS and ICBG	0.5, 1.5, 3, 6, 12, and 24 months
				
				% f/u NR
					• End plate resorption
					  • 82% (41/50) levels
					  • ALIF: 83% (10/12) levels
					  • PLIF: 100% (2/2) levels
					  • TLIF: 81% (29/36) levels
					  • Onset of resorption late compared to the cervical spine.
					  • Degree of resorption varied between patients and levels of patients who underwent more then 1 level of fusion.
					  • Transition to bone formation primarily occurred between 6-9 months.
					• Cage migration
					  • 28% (10/36) of patients
					  • ALIF: 10% (1/10) patients
					  • PLIF: 50% (1/2) patients
					  • TLIF: 33% (8/24) patients
					  • Occurred by 6 weeks
					  • Associated with, at re-exploration, an increase in the size of the intervertebral space.
					  • Responsible for neurologic symptoms in TLIF and PLIF patients only.
					  • Led to revision surgery in 8 patients.
					• Subsidence of disc space
					  • 22% (11/50) levels
					  • Mean disc space subsidence was 17.8 %

Villavicencio 2005	N = 74
	38% male
	57 Years
	
	<u>Doses</u>:
	4.2 or 12.0 mg/level
		DDD	1-3 level open or minimally invasive TLIF with and without posterolateral fusion with rhBMP-2/ACS, and local autograft and bone allograft.	3, 6, 12, and 24 months
				
				% f/u NR
					• Ectopic bone formation
					  • 0% (0/74)
					• Hematoma (clinical outcome not described)
					  • 3% (2/74)
					• Infection (clinical outcome not described)
					  • 3% (2/74)

Off-label use: rhBMP-7					

Furlan 2007 (OP-1 lumbar + cervical)	N = 30
	
	Lumbar fusions: N = 16
	
	Cervical fusions: N = 14
	
	43% female
	53 years
	
	<u>Doses</u>:
	7 mg/level (rhBMP-7)
		Patients at a high risk for pseudoarthrosis. These consisted of patients with connective tissue disorders, major medical comorbidities or medications that could interfere with bone healing, history of nonunion fusions, limited availability or poor quality of autogenous bone graft.	Posterolateral with rhBMP-7/ bovine type I collagen.	3, 6, 12, 18, and 24 months
				
				% f/u NR
					• Superficial wound infections
					  • 7% (2/30)
					  • Cervical and lumbar not reported separately.
					• Systemic toxicity
					  • 0%
					• Heterotopic ossification
					  • 0%
					• Epidural ossification
					  • 0%

Govender 2002	N = 9
	44% male
	47 years
	
	Doses:
	3.5 mg
		Stenosis, spodylolist-hesis, instability, chiari I malformati-on and basilar invagination, tethered cord syndrome, fracture	rhBMP-7 with collagen carrier and autogenous bone graft	Mean 5.22 months (2-15)
				
				% f/u NR
					• Myelopathy
					  • 11% (1/9) patients
					• Spondylosis
					  • 11% (1/9) patients
					• Spinal chord compression
					  • 11% (1/9)
					  • Required surgical intervention
					• Cerebrospinal fluid leak
					  • 11% (1/9)
					  • Required insertion of spinal drain

Vaccaro Patel 2003/2005 (pilot study)	N = 12
	25% male
	68 years
	
	Doses:
	7 mg/level (rhBMP-7)
		Degenerativ e lumbar spondylolist hesis with symptoms of neurogenic claudication.	1- level PLF with rhBMP-7/bovine type I collagen and ICBG	1.5, 3, 6, 9, 12, and 24 months.
				
				83% (10/12)
					• Ectopic bone formation:
					  • 0%
					• Local or systematic toxicity:
					  • 0%
					• Revision posterior lumbar fusion for pseudarthrosis
					  • 8% (1/12)

Anterior Cervical Discectomy and Fusion; ACS: Absorbable Collagen Sponge; ALIF: Anterior Lumbar Interbody Fusion; DDD: Degenerative Disc Disease; f/u: follow-up; ICBG: Iliac Crest Bone Graft; NR: data not reported; PEEK: Polytetheretherketone; PLIF: Posterior Lumbar Interbody Fusion; TLIF: Transforaminal Lumbar Interbody Fusion

*Demographic data reported only for the 146/277 patients with 6-year follow-up available; the authors stated that the demographic data was similar for the original group of 277 patients.
† Of the 83 patients that met the inclusion criteria, 50 consecutive patients completed a minimum of 12 months of clinical follow-up. ‡§§

Appendix Table 13. Case series evaluating the safety of BMPs in cervical spinal fusion.

Investigator (yr, country, ref #) Surgical Site	No. pts Sex Mean age (BMP dose)	Diagnosis	Surgical intervention	Follow-up: Duration Loss to f/u (%)	Reported complications
Off-label use: rhBMP-2					

Hamilton 2011	N = 60
	37% male
	56 years
	
	Doses:
	Mean : 1.8 mg/level
		Basilar invagination: 11% (6/53)
		
		Fracture: 11% (6/53)
		
		Atlantoaxial instability: 30% (16/53)
		
		Kyphosis/kyphoscoliosis: 41% (22/53)
		
		Osteomyelitis: 2% (1/53)
		
		Spondylolisthesis: 2% (1/53)
		
		Cyst: 2% (1/53)
			rhBMP-2/ACS with allograft or minimal autograft in some cases	40 months mean (25-80 months)
				
				88% f/u (53/60)
					• Neck swelling
					  • 0% (0/53) patients
					• Dysphagia
					  • 0% (0/53) patients
					• Superficial wound infection
					  • 2% (1/53)
					• Adjacent level degeneration
					  • 2% (1/53)

Hiremath 2009	N = 16
	19% male
	59 years
	
	Doses:
	0.75-4.05 mg/level.
	Mean = 1.95
		Failed ACDF, trauma, unhealed fracture, spondolytic myelopathy, rheumatoid arthritis or other (including neoplastic processes)	1-4 level posterior cervical or cervico-thoracic fusion with rhBMP-2/ACS allograft with additional graft material (ICBG local morselized bone, frafton Putty, or Vitoss), and instrumenta-tion.	3-14 months (mean = 5.7)
				
				% f/u NR
					• Neck swelling without hematoma
					  • 6% (1/16)
					  • Resolved with steroid treatment
					• Hematoma
					  • 0% (0/16)
					• Wound infection
					  • 0%
					• Dysphagia or other airway compromise
					  • 0%
					• Screw pullout
					  • 6% (1/16)
					  • Resulted in severe pain. The patient was not a candidate for reoperation due to comorbidities.
					• Broken rod
					  • 6% (1/16)
					  • Considered a minor failure and did not necessitate reoperation.

Klimo 2009	N = 22
	64% male
	53 years
	
	<u>Doses</u>:
	1.1-2.1 mg/level
		NR	rhBMP-2 in PEEK cage, anterior plate fixation	6,12,24, and 52 weeks
				
				0% (0/22) loss to f/u
					• Heterotopic bone formation*
					  • 32% (12/38) levels Grade 1 or 2
					  • 68% (26/36) levels Grade 3b
					• End plate resorption
					  • Classified as none, mild, moderate and severe.†
					  • 3 levels could not be assessed due to inadequate visualization
					  • 20% (7/35) levels had no resorption
					  • 23% (8/35) levels had mild resorption
					  • 57% (20/35) levels had moderate or severe resorption
					• Neck swelling
					  • 5% (1/22) patients
					  • Manifested on post-surgery day 2
					  • No airway compromise was noted and patient was discharged the next day
					• Recurrent laryngeal nerve palsy
					  • 5% (1/22) patients
					  • Occurred after three-level fusion
					  • Onset date NR
					  • Patient recovered after 3 months
					• Sphrerical radiolucencies
					  • 1-2mm – 6mm and larger in size
					  • 39% (15/38) levels
					  • Occurred in central core of PEEK grafts

Sethi 2011 (lumbar, cervical)	N = 95
	55% male
	51 years
	
	Lumbar:
	64% (61/95)
	87 levels
	
	Cervical:
	36% (34/95)
	50 levels
	
	<u>Doses</u>:
			Lumbar:
			ALIF: 38% (23/61)
			
			TLIF: 59% (36/61)
			
			PLIF: 3% (2/61)
			
			Cervical: Anterior cervical depression and fusion with stabilizing plate.
			
			Both lumbar and cervical:
			PEEK cage: 62% (59/95) or 60% (82/137) levels
			
			Allograft: 38% (36/95) or 40% (55/137) levels
				2 and 6 weeks, 3, 6, 12, and 24 months
				
				% f/u NR
					• End plate resorption (**cervical**)
					  • 100% (50/50) levels cervical spine
					  • Osteolysis of vertebral body in some patients
					  • Observed as early as 2 weeks post-op in some patients, and by 6 weeks all had experienced some form of resorption
					  • Largest transition occurred between 3 and 6 months
					• Subsidence/narrowing of disk space (**lumbar + cervical**)
					  • 50% of patients (47/95)
					  • Average subsidence for group was 16.5% at 12 months
					• Prevertebral swelling (**cervical**)
					  • 100% (34/34) patients
					  • Week 1: swelling measured 15.7 mm
					  • Week 2: swelling measured 11.8 mm
					  • Week 6: swelling measured 8.0 mm
					  • After 6 weeks swelling returned to near preoperative state
					• Cage migration (**lumbar + cervical**)
					  • Lumbar: 10-11/61 patients, with 10 of which underwent TLIF with a PEEK cage.
					  • TLIF with PEEK cage: 38% (10/36)
					  • Cervical: 0-1/34 patients
					  • Unclear if the one additional case occurred in lumbar or cervical.

Shen 2010	N = 127
	43% male
	54 years
	
	<u>Doses</u>:
	4 mg/level for 3-level fusion
	
	8 mg/level fir 4-and 5- level fusions
		Cervical spondylotic radiculopat-hy: 65% (83/127) patients
		
		Cervical spondylotic myelopathy or myeloradicul opathy: 35% (44/127) patients
			rhBMP-2 with structural allograft/ ACS: 83% (105/127)
			
			rhBMP-2 with PEEK cage/ACS: 8% (10/127)
			
			rhBMP-2 with titanium mesh cage/ACS: 9% (12/127)
			
			3-level fusion: 59% (75/127)
			
			4-level fusion: 27% (34/127)
			
			5-level fusion: 14.2% (18/127
			
			451 segments total.
				2 years minimum
				
				0% (0/127) loss to f/u
					• Revision surgery:
					  • 6.3% (8/127) (for pseudarthrosis)
					• Pseudoarthrosis
					  • 10% (13/127) of patients or 3% (14/451) levels of fusion
					  • Diagnosed at 6 months post-surgery
					  • 8 of these patients required revision surgery
					• Neck swelling and difficulty swallowing
					  • Reported in most cases, rate NR
					• Seroma
					  • 0% (0/127)
					• Hematoma
					  • 0% (0/127)

Shields 2006	N = 151
	41% male
	50 years
	
	Doses:
	≤2.1 mg/level
		Spondylosis: 74% (112/151)
		
		Herniation: 26% (39/151)
			1-3 level ACDF (N=138) or vertebrect-omy (N=13) with resorbable poly (D,C-lactic acid) cage or homologous bone graft filled with rhBMP-2	NR
				
				% f/u NR
					• Dysphagia, respiratory difficulties or incisional swelling (without hematoma)
					  • 9% (13/151)
					• Hematoma
					  • 10% (15/151)
					• Graft resorption
					  • 1% (1/151)
					• Implant dislodgement
					  • 1% (2/151)

Stachniak 2011	N = 30
	20% male
	53 years
	
	Doses:
	0.6 mg/level
		NR	ACDF with PEEK spacers, rhBMP-2/collagen sponge with titanium plates	2, 6, and 10 weeks; 6 months
				
				% f/u NR
					• Soft tissue swelling
					  • At base line, 93% (28/30) of patients had a mean swelling of 12.4 mm
					  • At 2 weeks, 70% (21/30) of patients had a mean swelling of 21.8 mm
					  • At 6 weeks, 80% (24/30) of patients had a mean swelling of 20.6 mm
					  • At 10 weeks 73% (22/30) of patients had a mean swelling of 18.4 mm
					  • At 6 months 70% (21/30) patients had a mean swelling of 14.2 mm
					• Dysphagia (SAW-QOL)
					  • At 2 weeks, 19% of patients frequently chocked on food
					  • At 2 weeks, 4.8% frequently chocked when drinking
					  • At 2 weeks, 48% experienced frequent food sticking in their throats.
					  • At 6 months, 0% of patients frequently chocked on food
					  • At 6 months, 6.7% of patients had difficulty drinking
					  • At 6 months, 6.7% of patients experienced frequent food sticking in their throats.

Tumialan 2008/Boakye 2005 (Tumialan includes all pts reported in Boakye 2005)	N = 200
	48% male
	54 years
	
	<u>Doses</u>:
	Initial 24 pts: 2.1 mg/level
	
	Next 93 pts: 1.05 mg/level
	
	Final 83 pts: 0.7 mg/level
		Myelopathy: 34% (68/200)
		
		Adjacent-segment disc herniations: 11% (22/200)
		
		Pseudoarthrosis: 5% (10/200)
		
		Non-specified: 50% (100/200)
			1-4 level ACDF with a rhBMP-2/ACS filled PEEK spacer.	17 months (8-26)
				
				% f/u NR
					• Dysphagia (presented post-peratively)
					  • 7% (14/200)
					  • Severe dysphagia
					  • 36% (5/14)
					  • Moderate dysphagia
					  • 21% (3/14)
					  • Mild dysphagia
					  • 43% (6/14)
					• Excess interbody bone formation
					  • 2% (3/200)
					  • Patients asymptomatic
					• Hematoma
					  • 1% (2/200)
					• Seroma
					  • 1% (2/200)

Vaidya 2008 (cervical + lumbar)	N = 59 (82 levels)
	
	Cervical fusions:
	N = 23 (32 levels)
	
	Lumbar fusions:
	N = 36 (50 levels)
	
	% female NR
	52 years
	
	<u>Doses</u>:
	1 mg/level
		Spondylolisthes-is, adult scoliosis, revision surgery, and discogenic pain	ACDF with rhBMP-2/ACS	0.5, 1.5, 3, 6, 12, and 24 months.
				
				% f/u NR
					• End plate resorption
					  • 100% (32/32) levels
					  • Detected by 2-6 weeks
					  • In all cases, occurred in both the superior and inferior end plates
					  • Earlier onset of resorption in comparison to the lumbar spine
					  • Transition to bone formation occurred between 3-6 months in the majority of cases
					• Cage migration
					  • 4% (1/23) patients
					  • Minimal and not associated with any clinical sequelae
					• Subsidence of disc space
					  • 41% (13/32) levels
					  • Mean disc space subsidence was 12.8%

Off-label use: rhBMP-7					

Furlan 2007 (OP-1 lumbar + cervical)	N = 30
	
	Cervical fusions: N = 14
	
	Lumbar fusions: N = 16
	
	43% female
	53 years
	
	<u>Doses</u>: 7mg/level rhBMP-7
		Patients at a high risk for pseudoarthrosis. These consisted of patients with connective tissue disorders, major medical comorbidities or medications that could interfere with bone healing, history of nonunion fusions, limited availability or poor quality of autogenous bone graft.	ACDF with rhBMP-7/ bovine type-I collag-en	3, 6, 12, 18, and 24 months
				
				% f/u NR
					• Superficial wound infections
					  • 7% (2/30)
					  • Not reported separately for lumbar versus cervical
					• Systemic toxicity
					  • 0%
					• Heterotopic ossification
					  • 7% (1/14)
					  • Asymptomatic
					• Peridural ossification
					  • 0%

Leach 2009	N = 131
	Sex NR
	Age NR
	
	Doses:
	1.75-3.5 mg/level
		NR	All patients had anterior interbody fusion using PEEK, carbon, or trabecular metal cages with or without an anterior cervical plate.
			
			rhBMP-7/ collagen with tricalcium phosphate: 94% (123/131) of patients.
			
			Tricalcium phosphate alone: 6% (8/131)
				Within the first 30 days
				
				0% (0/131) loss to f/u
					• Recurrent brachialgia (arm pain)
					  • 0.8% (1/131) patients
					  • 72 hours post-op
					• Dysphagia and dysphonia (sudden onset)
					  • 0.8% (1/131) patients
					  • 8 days post-op
					  • CT scan did not reveal unusual neck swelling or hematoma, no indication of laryngeal nerve dysfunction.
					  • Psychological factors may have resulted in her symptoms
					• Dysphagia (moderate)
					  • 0.8% (1/131) patients
					  • Occurred beyond 3 months post-op but resolved by 12 months
					  • No treatment required

ACDF: Anterior Cervical Discectomy and Fusion; ACS: Absorbable Collagen Sponge; ALIF: Anterior Lumbar Interbody Fusion; DDD: Degenerative Disc Disease; f/u: follow-up; NR: data not reported; PEEK: Polytetheretherketone; PLIF: Posterior Lumbar Interbody Fusion; SAW-QoL: Swallowing—Quality of Life evaluation; TLIF: Transforaminal Lumbar Interbody Fusion

* Classifed as Grade 1 (ossification in disc space exclusively), Grade 2 (ossification in to outer aspects of annulus), Grade 3a (ossification within spinal canal), Grade 3b (ossification within one foramen) and Grade 3c (ossification within both foramina).

† Mild resorption was defined as “minor indistinctness of endplates when compared with preoperative image”. Moderate resorption was defined as “more indistinctness of end plates when compared with preoperative image”. Severe resorption was defined as “ complete indistinctness of end plates when compared with preoperative image, difficulty in assessing where spinal body ends and interbody disc space begins.

Appendix Table 14. Case reports evaluating the safety of BMPs in lumbar spinal fusion.

Investigator (yr, country, ref #) Surgical Site	No. pts Sex Mean age (BMP dose)	Diagnosis	Surgical intervention	Duration follow-up	Reported complications
On-label (no case reports identified)					
Off-label use: rhBMP-2					

Balseiro 2010	N = 2
	100% male
	64 years
	
	Doses:
	4 mg rhBMP-2
		Disc-herniation, mechanical back pain in one case, postlaminect-omy instability in the other	L3-L5 TLIF with rhBMP-2/ACS filled PEEK cage.	15 months, 2 years	• Osteolysis
					  • 3 months post-op.
					  • Appeared to be result of preoperative subchondral cyst.
					  • Occurred in both cases.

Brower 2008	N = 1
	Male
	69 years
	
	Doses:
	12 mg/8.4 mL
		Degenerative disc disease, spondylolisth-esis, stenosis	L4-L5 laminectomy, intertransver-se fusion, rhBMP-2/ACS with pedicle screws	NR	• Heterotopic bone formation
					  • Appeared at 3 months
					  • Did not appear to have significant effect on recovery

Chen 2010	N = 4 50% male 61 years Doses:	DDD, spondylolisthe-sis, stenosis	Minimally invasive TLIF with rhBMP/2-ACS with rods and pedicle screws.	18 months for one patient, 12 for two and 63 months for the fourth.	• Radiculopathy due to ectopic bone growth appeared in all four cases.

Lastfogel 2010	N = 3 100% male 41 years Doses: NR	Spondylolisthe-sis	ALIF	1 year in two cases, 9 months in the third.	NR
Lewandrowski 2007	N = 5 40% female 50 years Doses: 4.2 mg	DDD	TLIF with rhBMP-2/ACS in PEEK cage	NR	• Vertebral osteolysis • Symptoms occurred between 4 weeks and 3 months post-op

Moshel 2008	N = 1
	Male
	53 years
	
	Dose: NR
		Back pain and radiculopathy	First operation: Capstone spacer with rhBMP-2
			
			Second operation: autologous bone graft without rhBMP
			
			Third operation: autograft with bovine collagen and rhBMP-2
				NR	• Transient supraventricular tachycardia
					  • Developed on post-op day 1 in the case of the first operation, and on day 10 after the third operation.
					• Sepsis
					  • Attributed to an immune response to the BMP

Muchow 2010	N = 1 27 years Male Doses:	DDD and stenosis	TLIF with rhBMP-2	NR	• Bone formation, surrounded by a fibrovascular stroma was discovered adjacent to the L4 nerve root. • The patient began complaining of pain at post-op week 4, but the mass was not discovered until fifteen weeks after the initial operation.

Steib 2010	N = 1 Male 23 years Doses: NR	Recurrent surgical malunions. The surgeries were initially undertaken to treat secondary hyperkyphosis, .	rhBMP-2 in an interbody cage.	NR	• Fatal neurofibromatosis occurring five months after operation.
Whang 2008	N = 1 Male 42 years Doses: 8.4 mg	"degenerative changes limited to the LF-S1 disk space"	rhBMP-2/ ACS, PEEK spacer and autogenous bone graft	6 weeks, 12 weeks and1 year after second surgery.	• Lack of an alleviation of symptoms necessitated a revision operation where an autograft was used.
Wong 2008	N = 5 40% male 31 years Doses:	Discogenic mechanical back pain, spondylolistes-is, radiculopathy,	PLIF = 20% (1/5) TLIF = 80% (4/5) Both used rhBMP-2/ACS	NR	• Ectopic bone growth in 100% (5/5) of the patients accompanied by radicular pain.

Off-label use: rhBMP-7					
Kim 2010	N = 1 Male 42 years Doses: 3 doses at 3.5 mg/dose. Total: 10.5 mg	Flat-back syndrome with symptomatic junctional degenerative disease	rhBMP-2/bovine collagen, local bone graft, autogenous bone graft, rod and pedicle screws.	3 and 10 months.	• Ectopic bone mass, removed at 10 months post-op.

Appendix Table 15. Case series evaluating the safety of BMPs in cervical spinal fusion.

Investigator (yr, country, ref #) Surgical Site	No. pts Sex Mean age (BMP dose)	Diagnosis	Surgical intervention	Follow-up: Duration Loss to f/u (%)	Reported complications
Off-label use: rhBMP-2					
Anderson Burton 2011	N = 2 100% male 56 years Doses:	Spondylosis, stenosis, pseudoarthrosis	ACDF with rhBMP-2		• Seroma • Appeared in one patient 2 weeks post-op and in the other patient 5 days post-op.
Perri 2007	N = 1 Male 54 years Doses:	NR	ACDF with rhBMP-2/ACS	NR	• Severe neck swelling
Robin 2010	N = 1 Female 66 years Doses: 2.1 mg	Spondylosis, stenosis	Bilateral laminectomy with posterolateral instrumentation and arthrodesis.	NR	• Seroma • Symptoms appeared on post-op day 5 • Tested positive for cytokines
Shahlaie 2008	N = 1 Female 53 years Doses: 12 mg	Basilar invagination with stenosis, spinal cord compression	rhBMP-2 with autograft	3 and 4 months	• Seroma • Removed on post-op day 3
Off-label use: rhBMP-7: no case reports identified					

Appendix Table 16. Case reports evaluating the safety of BMPs in thoracic spinal fusion.

Investigator (yr, country, ref #) Surgical Site	No. pts Sex Mean age (BMP dose)	Diagnosis	Surgical intervention	Duration follow-up	Reported complications
Off-label use: rhBMP-2					

Deutsch 2010	N = 1
	Male
	56 years
	
	Doses:
	12 mg/level
	
	then 6 mg/level to posterior lateral gutter
		Pseudoarthrosis and screw pullout from another operation.	Anterior interbody fusion with Grafton demineralized bone matrix, allograft, autogenous ribgraft and rhBMP-2	NR	• Ectopic bone formation
					  • Patient experienced 40 lbs weight loss, satiety and pain with urination
					  • Occurred over the first six months post-op
					• Seroma
					  • Appeared and drained at one month post-op.

Appendix Table 17. Differential efficacy or safety in various subpopulations.

Investigator	Outcomes mean ± SD (unless otherwise indicated) (range)
Slosar (2007) prospective cohort **Lumbar spine** Treatment groups: *rhBMP2*: n = 45 *autograft*: n = 30	**Radiographic Outcomes**
	Non-union rate based on X-ray and CT scan by levels treated (f/u period NR) (n, % patients with non-union) ***1-level vs. 2-level vs. 3-level*** rhBMP2: 0% (0/10) vs. 0% (0/26) vs. 0% (0/9) autograft: 11% (1/9) vs. 13% (2/15) vs. 33% (2/6)
Glassman, Carreon (2007) * Retrospective cohort with historical control **Lumbar spine** Treatment groups: *rhBMP2*: n = 91 *ICBG*: n = 35	**Radiographic Outcomes**
	Non-union rate based on fine-cut CT scan (n, % patients with non-union) ***Males vs. females*** rhBMP2: 11.1% (4/36) vs. 3.6% (2/55) ICBG: 26% (NR) vs. 0% (NR) ***Smokers vs. non-smokers*** rhBMP2: 0% (0/14) vs. 7.8% (6/77) ICBG: 40% (2/5) vs. 10% (3/30)
	CT grade based on fine-cut CT scan (mean) ***Males vs. females*** rhBMP2: 4.04 vs. 4.61 ICBG: 3.75 vs. 4.69 ***Smokers vs. non-smokers*** rhBMP2: 4.32 vs. 4.40 ICBG: 3.20 vs. 4.33
Glassman, Dimar (2007) † Retrospective cohort with historical control **Lumbar spine** Treatment groups: *rhBMP2*: n = 76 *ICBG*: n = 72	**Radiographic Outcomes**
	Fusion rate based on IDE fusion success criteria (% patients with fusion) ***Smokers vs. Non-smokers at 12 month f/u*** rhBMP2: 94.7% vs. 96.3% ICBG: 75.0% vs. 89.6% ***Smokers vs. Non-smokers at 24 month f/u*** rhBMP2: 95.2% (20/21) vs. 100.0% (55/55) ICBG: 76.2% (16/21) vs. 94.1% (48/51) **Fusion rate based on CT scan bridging bone criteria (% patients with fusion)** ***Smokers vs. Non-smokers at 12 month f/u*** rhBMP2: 94.4% vs. 94.4% ICBG: 73.7% vs. 83.3% ***Smokers vs. Non-smokers at 24 month f/u*** rhBMP2: 95.0% (19/20) vs. 98.1% (52/53) ICBG: 75.0% (15/20) vs. 90.2% (46/51)
	Pain
	Improvement in ODI score from pre-operative score (mean) ***Smokers vs. Non-smokers at 24 month f/u*** rhBMP2: 22.1 vs. 26.4 ICBG: 21.0 vs. 24.6

Investigator	Outcomes mean ± SD (unless otherwise indicated) (range)
	Function
	Improvement in SF-36 PCS score from pre-operative score (mean) ***Smokers vs. Non-smokers at 24 month f/u*** rhBMP2: 7.1 vs. 10.2 ICBG: 11.6 vs. 11.2
Cahill et al. (2009) Retrospective cohort (database) study **Cervical spine** (subset of total population) Treatment groups: ***Anterior cervical*** *rhBMP (any)*: n = 2299 *Non-BMP*: n = 24,768 ***Posterior cervical*** *rhBMP (any)*: n = 478 *Non-BMP*: n = 2391	**Surgical and perioperative complications**
	Overall complication rate (n, % patients) ***Anterior cervical vs. Posterior cervical fusion*** fusion with rhBMP (any): 7.09% (163/2299) vs. 10.04% (48/478) fusion without rhBMP: 4.68% (1158/24768) vs. 9.95% (238/2391) **Dysphagia or hoarseness rate (n, % patients)** ***Anterior cervical vs. Posterior cervical fusion*** fusion with rhBMP (any): 4.35% (100/2299) vs. 2.09% (10/478) fusion without rhBMP: 2.45% (608/24768) vs. 1.63% (39/2391) **Wound complication rate (n, % patients)** ***Anterior cervical vs. Posterior cervical fusion*** fusion with rhBMP (any): 1.22% (28/2299) vs. 2.93% (14/478) fusion without rhBMP: 0.65% (160/24768) vs. 2.51% (60/2391)
Taghavi (2010) retrospective cohort **Lumbar spine** Treatment groups: *rhBMP2*: n = 24 *BMAA*: n = 18 *autograft*: n = 20	**Radiographic Outcomes**
	Time to solid fusion (days) ***1-level vs. multi-level*** rhBMP2: 199.8 ± 49.8 vs. 240.4 ± 71.3 BMAA: 313.3 ± 34.3 vs. 282.0 ± 87.5 autograft: 276.7 ± 29.8 vs. 263.3 ± 79.4 **Fusion rate (% patients with fusion)** ***1-level vs. multi-level*** rhBMP2: 100% (13/13) vs. 100% (11/11) BMAA: 100% (7/7) vs. 63.6% (7/11) autograft: 100% (10/10) vs. 100% (10/10)
Carragee, Mitsunaga (2011) Retrospective cohort **Lumbar spine** Treatment groups: *rhBMP2*: n = 69 *no rhBMP2*: n = 72	**Adverse events**
	RE complication rate (n, % patients, 90% CI) ***1-level vs. 2-level fusion*** fusion with rhBMP2: 6.7% (3/45, 0.55 – 12.79) vs. 8.3% (2/24, -0.95 – 17.61) fusion without rhBMP2: 0% (0/110, < 2.4) vs. 1.6% (1/64, -0.99 – 4.11)
Deyo et al. (2011) ‡	**Second Surgeries**

Investigator	Outcomes mean ± SD (unless otherwise indicated) (range)
Retrospective cohort (database) study **Lumbar spine** Treatment groups: *rhBMP (any)*: n = 1703 *Non-BMP*: n = 15,119	**Repeat surgery within 1 year of index surgery (n, % patients)** ***Previous surgery vs. no previous surgery*** rhBMP (any): 3.8% (14/366) vs. 2.4% (32/1337) non-BMP: 4.6% (100/2181) vs. 2.7% (343/12938) ***Simple fusion vs. Complex fusion*** rhBMP (any): 2.6% (26/1014) vs. 2.9% (20/689) non-BMP: 2.8% (307/10792) vs. 3.1% (136/4327) **Repeat surgery within 2 years of index surgery (n, % patients)** ***Previous surgery vs. no previous surgery*** rhBMP (any): 8.2% (30/366) vs. 5.8% (77/1337) non-BMP: 8.5% (186/2181) vs. 5.6% (726/12938) ***Simple fusion vs. Complex fusion*** rhBMP (any): 6.1% (62/1014) vs. 6.5% (45/689) non-BMP: 5.8% (630/10792) vs. 6.5% (282/4327) **Repeat surgery within 3 years of index surgery (n, % patients)** ***Previous surgery vs. no previous surgery*** rhBMP (any): 12.3% (45/366) vs. 8.4% (112/1337) non-BMP: 12.1% (264/2181) vs. 7.9% (1023/12938) ***Simple fusion vs. Complex fusion*** rhBMP (any): 8.9% (90/1014) vs. 9.7% (67/689) non-BMP: 8.2% (881/10792) vs. 9.4% (406/4327) **Repeat surgery within 4 years of index surgery (n, % patients)** ***Previous surgery vs. no previous surgery*** rhBMP (any): 14.5% (53/366) vs. 9.7% (130/1337) non-BMP: 14.9% (325/2181) vs. 9.8% (1263/12938) ***Simple fusion vs. Complex fusion*** rhBMP (any): 10.0% (101/1014) vs. % 11.9 (82/689) non-BMP: 10.3% (1092/10792) vs. 11.5% (496/4327)
Williams (2011)[51§] Retrospective cohort (database) study *rhBMP (any)*: n = 11,933 *Non-BMP*: n = 43,929	**Surgical and perioperative complications** **Overall complication rate (n, % patients)** ***Adult scoliosis vs. Pediatric scoliosis*** fusion with rhBMP (any): 13.8% (124/899) vs. 8.8% (139/1576) fusion without rhBMP: 9.3% (425/4586) vs. 7.0% (1310/15937) **Superficial infection (n, % patients)** ***Adult scoliosis vs. Pediatric scoliosis*** fusion with rhBMP (any): 1.3% (12/899) vs. 1.1% (18/1576) fusion without rhBMP: 0.9% (42/4586) vs. 0.7% (138/15937) **Deep infection (n, % patients)** ***Adult scoliosis vs. Pediatric scoliosis*** fusion with rhBMP (any): 1.8% (16/899) vs. 1.6% (26/1576) fusion without rhBMP: 2.0% (90/4586) vs. 1.3% (235/15937) **Epidural hematoma/seroma (n, % patients)** ***Adult scoliosis vs. Pediatric scoliosis*** fusion with rhBMP (any): 0.1% (1/899) vs. 0.2% (3/1576) fusion without rhBMP: 0.3% (13/4586) vs. 0.1% (20/15937)

f/u: follow-up; SD: standard deviation; IDE: Investigational Device Exemption; ODI: Oswestry Disability Index; SF-36: Short-Form 36; PCS: Physical Component Summary; RE: Retrograde Ejaculation; CI: Confidence Interval; BMAA: Bone Marrow Aspirate with Allograft

* rhBMP2 group is a mixture of one-level (n = 61) and two-level (n=30) treatments, ICBG control group is one-level treatment only (n = 35). CT grade based on the following criteria: grade 1 (no fusion) and grade 2

(partial unilateral fusion) defined as non-union; grade 3 (partial bilateral fusion) defined as probably fusion; grades 4 and 5 (solid unilateral or bilateral fusion) defined as definite fusion[23].

†Fusion success is defined by the IDE protocol as bilateral bridging trabecular bone on plain radiographs with less than 3° of translation and less than 5° of angulation on flexion-extension views; defined by CT scan criteria as presence of contiguous bridging bone on fine cut CT scan with coronal and sagittal reconstructions[62].

‡Previous surgery is defined as having had lumbar surgery prior to the index operation; repeat surgery is defined as any reoccurrence of lumbar surgery following the index operation, with the nature of surgery and spinal levels unknown. Simple fusion is defined as anterior fusion, transverse process or posterior fusion involving one or two disc levels, or an unreported number of disc levels; complex fusion is defined as 360-degree spine fusion by single incision, any combination of anterior with either transverse process or posterior fusion, or any fusion involving more than two disc levels[49].

§Authors focused on intraoperative and immediate postoperative complications, including death, new neurological deficit, wound infection (superficial or deep), pulmonary embolus, deep venous thrombosis, other pulmonary complications, implant related, peripheral nerve deficit, visual deficit, and epidural hematoma. Epidural hematoma and seroma complications are grouped together as "epidural hematoma/seroma". Scoliosis patients are separated into adult (≥ 21 years) and pediatric[51].

Appendix Table 18. Detailed results from studies evaluating the cost effectiveness of BMPs in the spine.

Study (year) country	Study design	Model details/ assumptions	Sensitivity analysis	Relevant results	Author conclusions
AHRQ (2010), United States	Cost-effectiveness analysis, based on Burkus 2002 RCT[14] (on-label rhBMP2 vs ICBG)	Payer (CMS) perspective. Stationary Markov models used, with three health states for the treatment group (prefusion, secondary intervention, and fusion) and six for the control group (same as above, with or without donor site pain). Minimum time to both union and fusion assumed to be six weeks	One-way, selected two-and three way analyses	Base case (BMP cost bundled into Medicare DRG payment): BMP dominant over ICBG: cost savings of $94, increase in 0.024 QALYs over 24 months. One-way sensitivity analyses of base case (various): BMP dominant treatment strategy in all but one analysis (cost savings ranging from $15-%1130 and increase in QALY from 0.018-0.051). BMP as added cost BMP no longer the dominant treatment strategy (assumes additional cost of $3000).	"Bundling the BMP cost into the Medicare DRG payment results in almost identical costs for treatment and control groups, thus rhBMP should be the dominant strategy. However, analyses that assume added rhBMP costs may reflect the more common payer strategy. Given the analyses that examine rhBMP as an added cost, the group treated with rhBMP had higher QALYs and higher costs."
Garrison (2007), United King-dom	Cost-effectiveness analysis, based on Burkus 2002 RCT[14] (on-label rhBMP2 vs ICBG)	Modified ABACUS economic model (developed by ABACUS International, model development funded by Medtronic)	None	rhBMP2 use increases cost to UK NHS by £1.3 million per year (adjusted) compared to cost of ICBG. Estimated incremental cost per QALY gained= £120,390. Probability that rhBMP2 is cost-effectiveness at willingness to pay threshold of £30,000 per QALY = 6.4%	"Use of BMP for spinal fusion is unlikely to be cost-effective"
Carreon (2009), United States	Cost-utility analysis using data from own RCT	Decision tree	None	Mean total two year cost = $2295 more for rhBMP2 vs. ICBG	"In patients over 60 years old, the use of rhBMP2/ACS was more cost-effective than

	(off-label rhBMP2 vs ICBG in patients ≥ 60 years of age; single- or mutli-level fusion).			Decision tree analysis results: Cost of using rhBMP2 = $39,967 with 0.11 mean improvement in the SF-6C; the cost of using ICBG = $42,286 with a mean improvement of 0.10 in SF-6D.	ICBG for posterolateral fusion."

Appendix G. CLINICAL PEER REVIEWERS

Reviewer
Drew Brian, M.D.
Michael Lee, M.D.

1. Goldberg CJ, Moore DP, Fogarty EE, Dowling FE. Scoliosis: a review. Pediatric surgery international 2008;24:129-44.
2. Hadjidakis DJ, Androulakis, II. Bone remodeling. Annals of the New York Academy of Sciences 2006;1092:385-96.
3. Hu SS, Tribus CB, Diab M, Ghanayem AJ. Spondylolisthesis and spondylolysis. J Bone Joint Surg Am 2008;90:656-71.
4. Kendirci M, Hellstrom W. Retrograde ejaculation: Etiology, diagnosis, and management. Current Sexual Health Reports 2006;3:133-8.
5. Phillips B, Ball C, Sakett D, et al. Levels of evidence and grades of recommendation. 2001. (Accessed December, 2006, at http://www.cebm.net/levels_of_evidence.asp.)
6. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. BMJ 2004;328:1490.
7. West S, King V, Carey TS, et a. Systems to Rate the Strength of Scientific Evidence. Evidence Report/Technology Assessment No. 47 (Prepared by the Research Triangle Institute-University of North Carolina Evidence-based Practice Center, Contract No. 290-97-0011): Agency for Healthcare Research and Quality. Rockville, MD; 2002.
8. West S, King V, Carey TS, et al. Systems to rate the strength of scientific evidence. Evid Rep Technol Assess (Summ) 2002:1-11.
9. Ofman JJ, Sullivan SD, Neumann PJ, et al. Examining the value and quality of health economic analyses: implications of utilizing the QHES. J Manag Care Pharm 2003;9:53-61.
10. Chiou CF, Hay JW, Wallace JF, et al. Development and validation of a grading system for the quality of cost-effectiveness studies. Med Care 2003;41:32-44.
11. Karppinen J, Malmivaara A, Kurunlahti M, et al. Periradicular infiltration for sciatica: a randomized controlled trial. Spine (Phila Pa 1976) 2001;26:1059-67.
12. Karppinen J, Ohinmaa A, Malmivaara A, et al. Cost effectiveness of periradicular infiltration for sciatica: subgroup analysis of a randomized controlled trial. Spine (Phila Pa 1976) 2001;26:2587-95.
13. Boden SD, Zdeblick TA, Sandhu HS, Heim SE. The use of rhBMP-2 in interbody fusion cages. Definitive evidence of osteoinduction in humans: a preliminary report. Spine 2000;25:376-81.
14. Burkus JK, Gornet MF, Dickman CA, Zdeblick TA. Anterior lumbar interbody fusion using rhBMP-2 with tapered interbody cages. J Spinal Disord Tech 2002;15:337-49.
15. Burkus JK, Heim SE, Gornet MF, Zdeblick TA. Is INFUSE bone graft superior to autograft bone? An integrated analysis of clinical trials using the LT-CAGE lumbar tapered fusion device. J Spinal Disord Tech 2003;16:113-22.
16. Boden SD, Kang J, Sandhu H, Heller JG. Use of recombinant human bone morphogenetic protein-2 to achieve posterolateral lumbar spine fusion in humans: a prospective, randomized clinical pilot trial: 2002 Volvo Award in clinical studies. Spine 2002;27:2662-73.

17. Burkus JK, Sandhu HS, Gornet MF, Longley MC. Use of rhBMP-2 in combination with structural cortical allografts: clinical and radiographic outcomes in anterior lumbar spinal surgery. J Bone Joint Surg Am 2005;87:1205-12.
18. Burkus JK, Sandhu HS, Gornet MF. Influence of rhBMP-2 on the healing patterns associated with allograft interbody constructs in comparison with autograft. Spine 2006;31:775-81.
19. Dawson E, Bae HW, Burkus JK, Stambough JL, Glassman SD. Recombinant human bone morphogenetic protein-2 on an absorbable collagen sponge with an osteoconductive bulking agent in posterolateral arthrodesis with instrumentation. A prospective randomized trial. J Bone Joint Surg Am 2009;91:1604-13.
20. Dimar JR, 2nd, Glassman SD, Burkus JK, Pryor PW, Hardacker JW, Carreon LY. Clinical and radiographic analysis of an optimized rhBMP-2 formulation as an autograft replacement in posterolateral lumbar spine arthrodesis. J Bone Joint Surg Am 2009;91:1377-86.
21. Glassman SD, Carreon LY, Djurasovic M, et al. RhBMP-2 versus iliac crest bone graft for lumbar spine fusion: a randomized, controlled trial in patients over sixty years of age. Spine 2008;33:2843-9.
22. Haid RW, Jr., Branch CL, Jr., Alexander JT, Burkus JK. Posterior lumbar interbody fusion using recombinant human bone morphogenetic protein type 2 with cylindrical interbody cages. Spine J 2004;4:527-38; discussion 38-9.
23. Glassman SD, Carreon L, Djurasovic M, et al. Posterolateral lumbar spine fusion with INFUSE bone graft. Spine J 2007;7:44-9.
24. Mummaneni PV, Pan J, Haid RW, Rodts GE. Contribution of recombinant human bone morphogenetic protein-2 to the rapid creation of interbody fusion when used in transforaminal lumbar interbody fusion: a preliminary report. Invited submission from the Joint Section Meeting on Disorders of the Spine and Peripheral Nerves, March 2004. J Neurosurg Spine 2004;1:19-23.
25. Pradhan BB, Bae HW, Dawson EG, Patel VV, Delamarter RB. Graft resorption with the use of bone morphogenetic protein: lessons from anterior lumbar interbody fusion using femoral ring allografts and recombinant human bone morphogenetic protein-2. Spine 2006;31:E277-84.
26. Singh K, Smucker JD, Gill S, Boden SD. Use of recombinant human bone morphogenetic protein-2 as an adjunct in posterolateral lumbar spine fusion: a prospective CT-scan analysis at one and two years. J Spinal Disord Tech 2006;19:416-23.
27. Slosar PJ, Josey R, Reynolds J. Accelerating lumbar fusions by combining rhBMP-2 with allograft bone: a prospective analysis of interbody fusion rates and clinical outcomes. Spine J 2007;7:301-7.
28. Glassman SD, Dimar JR, 3rd, Burkus K, et al. The efficacy of rhBMP-2 for posterolateral lumbar fusion in smokers. Spine (Phila Pa 1976) 2007;32:1693-8.
29. Carragee EJ, Mitsunaga KA, Hurwitz EL, Scuderi GJ. Retrograde ejaculation after anterior lumbar interbody fusion using rhBMP-2: a cohort controlled study. Spine J 2011;11:511-6.
30. Crawford CH, 3rd, Bridwell KH, Cho W, et al. Extension of prior idiopathic scoliosis fusions to the sacrum: a matched cohort analysis of sixty patients with minimum two-year follow-up. Spine (Phila Pa 1976) 2010;35:1843-8.

31. Howard JM, Glassman SD, Carreon LY. Posterior iliac crest pain after posterolateral fusion with or without iliac crest graft harvest. Spine J 2011;11:534-7.
32. Joseph V, Rampersaud YR. Heterotopic bone formation with the use of rhBMP2 in posterior minimal access interbody fusion: a CT analysis. Spine (Phila Pa 1976) 2007;32:2885-90.
33. 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:E231-7.
34. 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;19:924-30.
35. Rihn JA, Patel R, Makda J, et al. Complications associated with single-level transforaminal lumbar interbody fusion. Spine J 2009;9:623-9.
36. Taghavi CE, Lee KB, Keorochana G, Tzeng ST, Yoo JH, Wang JC. Bone morphogenetic protein-2 and bone marrow aspirate with allograft as alternatives to autograft in instrumented revision posterolateral lumbar spinal fusion: a minimum two-year follow-up study. Spine (Phila Pa 1976) 2010;35:1144-50.
37. Vaidya R, Weir R, Sethi A, Meisterling S, Hakeos W, Wybo CD. Interbody fusion with allograft and rhBMP-2 leads to consistent fusion but early subsidence. J Bone Joint Surg Br 2007;89:342-5.
38. Burkus JK, Gornet MF, Glassman SD, et al. Blood Serum Antibody Analysis and Long -Term Follow-up of Patients Treated with Recombinant Human Bone Morphogenetic Protein-2 in the Lumbar Spine. Spine (Phila Pa 1976) 2011.
39. Vaccaro AR, Patel T, Fischgrund J, et al. A pilot study evaluating the safety and efficacy of OP-1 Putty (rhBMP-7) as a replacement for iliac crest autograft in posterolateral lumbar arthrodesis for degenerative spondylolisthesis. Spine (Phila Pa 1976) 2004;29:1885-92.
40. Vaccaro AR, Anderson DG, Patel T, et al. Comparison of OP-1 Putty (rhBMP-7) to iliac crest autograft for posterolateral lumbar arthrodesis: a minimum 2-year follow-up pilot study. Spine (Phila Pa 1976) 2005;30:2709-16.
41. Vaccaro AR, Whang PG, Patel T, et al. The safety and efficacy of OP-1 (rhBMP-7) as a replacement for iliac crest autograft for posterolateral lumbar arthrodesis: minimum 4-year follow-up of a pilot study. Spine J 2008;8:457-65.
42. Johnsson R, Stromqvist B, Aspenberg P. Randomized radiostereometric study comparing osteogenic protein-1 (BMP-7) and autograft bone in human noninstrumented posterolateral lumbar fusion: 2002 Volvo Award in clinical studies. Spine 2002;27:2654-61.
43. Kanayama M, Hashimoto T, Shigenobu K, Yamane S, Bauer TW, Togawa D. A prospective randomized study of posterolateral lumbar fusion using osteogenic protein-1 (OP-1) versus local autograft with ceramic bone substitute: emphasis of surgical exploration and histologic assessment. Spine 2006;31:1067-74.
44. Vaccaro AR, Lawrence JP, Patel T, et al. The safety and efficacy of OP-1 (rhBMP-7) as a replacement for iliac crest autograft in posterolateral lumbar arthrodesis: a long-term (>4 years) pivotal study. Spine 2008;33:2850-62.

45. Hwang CJ, Vaccaro AR, Hong J, et al. Immunogenicity of osteogenic protein 1: results from a prospective, randomized, controlled, multicenter pivotal study of uninstrumented lumbar posterolateral fusion. J Neurosurg Spine 2010;13:484-93.
46. Delawi D, Dhert WJ, Rillardon L, et al. A prospective, randomized, controlled, multicenter study of osteogenic protein-1 in instrumented posterolateral fusions: report on safety and feasibility. Spine (Phila Pa 1976) 2010;35:1185-91.
47. 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;302:58-66.
48. Cahill KS, Chi JH, Groff MW, McGuire K, Afendulis CC, Claus EB. Outcomes for Single-Level Lumbar Fusion: The Role of Bone Morphogenetic Protein. Spine (Phila Pa 1976) 2011.
49. Deyo RA, Ching A, Matsen L, et al. Use of Bone Morphogenetic Proteins in Spinal Fusion Surgery for Older Adults with Lumbar Stenosis: Trends, Complications, Repeat Surgery, and Charges. Spine (Phila Pa 1976) 2011.
50. 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;20:111-8.
51. Williams BJ, Smith JS, Fu KM, et al. Does BMP increase the incidence of perioperative complications in spinal fusion? A comparison of 55,862 cases of spinal fusion with and without BMP. Spine (Phila Pa 1976) 2011.
52. Baskin DS, Ryan P, Sonntag V, Westmark R, Widmayer MA. A prospective, randomized, controlled cervical fusion study using recombinant human bone morphogenetic protein-2 with the CORNERSTONE-SR allograft ring and the ATLANTIS anterior cervical plate. Spine 2003;28:1219-24; discussion 25.
53. Buttermann GR. Prospective nonrandomized comparison of an allograft with bone morphogenic protein versus an iliac-crest autograft in anterior cervical discectomy and fusion. Spine J 2008;8:426-35.
54. Crawford CH, 3rd, Carreon LY, McGinnis MD, Campbell MJ, Glassman SD. Perioperative complications of recombinant human bone morphogenetic protein-2 on an absorbable collagen sponge versus iliac crest bone graft for posterior cervical arthrodesis. Spine 2009;34:1390-4.
55. Smucker JD, Rhee JM, Singh K, Yoon ST, Heller JG. Increased swelling complications associated with off-label usage of rhBMP-2 in the anterior cervical spine. Spine 2006;31:2813-9.
56. 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;16:1257-65.
57. Xu R, Bydon M, Sciubba DM, et al. Safety and efficacy of rhBMP2 in posterior cervical spinal fusion for subaxial degenerative spine disease: Analysis of outcomes in 204 patients. Surg Neurol Int 2011;2:109.
58. Yaremchuk KL, Toma MS, Somers ML, Peterson E. Acute airway obstruction in cervical spinal procedures with bone morphogenetic proteins. Laryngoscope 2010;120:1954-7.

59. Burkus JK, Dorchak JD, Sanders DL. Radiographic assessment of interbody fusion using recombinant human bone morphogenetic protein type 2. Spine 2003;28:372-7.
60. Maeda T, Buchowski JM, Kim YJ, Mishiro T, Bridwell KH. Long adult spinal deformity fusion to the sacrum using rhBMP-2 versus autogenous iliac crest bone graft. Spine (Phila Pa 1976) 2009;34:2205-12.
61. FDA (US. Food and Drug Administration). Summary of safety and effectiveness (SSED): P000058: InFUSE Bone Graft/LT-CAGE Lumbar Tapered Fusion Device (Medtronic Sofamor Danek; Memphis TN); www.accessdata.fda.gov/cdrh_docs/pdf/P000058b.pdf (accessed August, 2011). 2002.
62. Garrett MP, Kakarla UK, Porter RW, Sonntag VK. Formation of painful seroma and edema after the use of recombinant human bone morphogenetic protein-2 in posterolateral lumbar spine fusions. Neurosurgery 2010;66:1044-9; discussion 9.