
DATE: 08 March 2013

CONTEXT AND POLICY ISSUES

Bone morphogenetic proteins (BMPs) are a group of growth factors capable of inducing the formation of bone. In addition to other vital tissue development functions, these proteins signal new bone growth through multiple pathways during the healing of bone. Osteoblasts are one cell type that respond to BMP signaling and are directly responsible for deposition of new bone. BMPs induce the differentiation of osteoprogenitor cells into osteoblasts and stimulate migration of these cells to areas of repair leading to the formation of new bone.\(^1\)

Formation of new bone is critical for surgical success in spinal fusion surgeries and in the treatment of nonunion of fractured limbs. In order to achieve sufficient production of new bone for these procedures a bone graft can be used to induce new bone formation. More recently, application of recombinant human BMPs has been an available alternative or adjunct to bone graft procedures in order to achieve sufficient new bone formation.\(^1\)

Two recombinant human BMPs, recombinant human BMP-2 (rhBMP-2) and recombinant human BMP-7 (rhBMP-7) are commercially available. RhBMP-2 is marketed as INFUSE (Medtronic, Minneapolis) and is approved for spinal fusion surgery using another product, the LT-CAGE. RhBMP-7 is marketed as OP-1 (Stryker Biotech, Hopkinton, Mass.) and is approved for use following nonunion of long bone fractures. Both products were approved for the described use by Health Canada in 2002.\(^2\)

As with many growth factors, disruption of normal BMP signaling is routinely found in tumour cells.\(^3\) Therefore at therapeutic doses the potential of exogenous BMPs to increase cancer incidence is a logical concern. Additional potential adverse events specific to these biotherapeutics must also be evaluated before their use is considered in any application.

Disclaimer: The Rapid Response Service is an information service for those involved in planning and providing health care in Canada. Rapid responses are based on a limited literature search and are not comprehensive, systematic reviews. The intent is to provide a list of sources of the best evidence on the topic that CADTH could identify using all reasonable efforts within the time allowed. Rapid responses should be considered along with other types of information and health care considerations. The information included in this response is not intended to replace professional medical advice, nor should it be construed as a recommendation for or against the use of a particular health technology. Readers are also cautioned that a lack of good quality evidence does not necessarily mean a lack of effectiveness particularly in the case of new and emerging health technologies, for which little information can be found, but which may in future prove to be effective. While CADTH has taken care in the preparation of the report to ensure that its contents are accurate, complete and up to date, CADTH does not make any guarantee to that effect. CADTH is not liable for any loss or damages resulting from use of the information in the report.

Copyright: This report contains CADTH copyright material and may contain material in which a third party owns copyright. This report may be used for the purposes of research or private study only. It may not be copied, posted on a web site, redistributed by email or stored on an electronic system without the prior written permission of CADTH or applicable copyright owner.

Links: This report may contain links to other information available on the websites of third parties on the Internet. CADTH does not have control over the content of such sites. Use of third party sites is governed by the owners’ own terms and conditions.
The purpose of this report is to retrieve and review the existing evidence on clinical effectiveness and safety of BMPs for use in spinal fusion surgery and treatment of long bone trauma.

**RESEARCH QUESTIONS**

1. What is the evidence for the clinical effectiveness of bone morphogenetic proteins used in spinal surgery or surgery for long bone trauma?

2. What is the evidence for the safety of bone morphogenetic proteins used in spinal surgery or surgery for long bone trauma?

**KEY FINDINGS**

The evidence presented in this report supports the use of rhBMP-2 in posterolateral lumbar fusion (PLF) and posterior lumbar interbody fusion (PLIF) as some outcomes may be superior to that of autograft alone. There are no statistically significant differences in clinical effectiveness outcomes for rhBMP-2 use in the treatment of long bone nonunion yet there is some limited evidence that rhBMP-2 use has the possible advantages of significantly less treatment failure and fewer secondary interventions for treatment of tibial fractures. Although the evidence reviewed in this report does not directly compare rhBMP-2 with rhBMP-7, it does provide more support for the use of rhBMP-2 in spinal fusion surgeries and the treatment of long bone nonunion. There is limited evidence to suggest serious acute adverse events following the use of BMPs in cervical spinal fusion surgeries. There was limited evidence of a significant increase in the incidence of retrograde ejaculation (RE) following anterior lumbar interbody fusion (ALIF) surgery using BMPs. This report also finds a significant evidence gap regarding an association between cancer incidence and BMP exposure.

**METHODS**

**Literature Search Strategy**

A limited literature search was conducted on key resources including PubMed, The Cochrane Library (2013, Issue 1), University of York Centre for Reviews and Dissemination (CRD) databases, Canadian and major international health technology agencies, as well as a focused Internet search. No filters were applied to limit the retrieval by study type. Where possible, retrieval was limited to the human population. The search was also limited to English language documents published between January 1, 2008 and February 6, 2013.

**Selection Criteria and Methods**

Table 1: Selection Criteria

| Population        | Spinal surgery patients
<table>
<thead>
<tr>
<th></th>
<th>Orthopedic surgery patients</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intervention</strong></td>
<td>Surgery with application of bone morphogenetic proteins (BMP) without autograft</td>
</tr>
<tr>
<td><strong>Comparator</strong></td>
<td>Autograft and surgery without application of bone morphogenetic proteins</td>
</tr>
<tr>
<td><strong>Outcomes</strong></td>
<td>Clinical effectiveness – bone fusion</td>
</tr>
<tr>
<td></td>
<td>Complications and safety</td>
</tr>
</tbody>
</table>
**Study Designs**

<table>
<thead>
<tr>
<th>Study Designs</th>
<th>Systematic reviews/meta-analyses, health technology assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Randomized controlled trials</td>
</tr>
<tr>
<td></td>
<td>Non-randomized studies</td>
</tr>
</tbody>
</table>

**Exclusion Criteria**

Studies were excluded if they did not meet the selection criteria, were duplicate publications, were included in a systematic review, or were published prior to 2008.

**Critical Appraisal of Individual Studies**

The quality of included systematic reviews was analyzed using the Assessment of Multiple Systematic Reviews (AMSTAR) tool. The quality of the included studies was assessed using the Downs and Black checklist. Instead of assignment of numerical scores for each study, strengths and limitations were described narratively and tabulated in Appendix 3.

**SUMMARY OF EVIDENCE**

**Quantity of Research Available**

The literature search strategy identified 363 articles and one additional article was found through searching grey literature. Full text retrieval of 56 articles followed screening of titles and available abstracts. Upon review, six systematic reviews, two randomized controlled trials (RCTs), four nonrandomized prospective controlled studies (PCSs) and twelve retrospective cohort studies (RCSs) met the selection criteria (Table 1). The excluded 32 studies consisted of twelve that were included in systematic reviews, five were reviews or letters, one was retracted, one was an English summary of a non-English publication and the remaining 13 were irrelevant and did not meet the inclusion criteria. A PRISMA flowchart describes the selection procedure of the included studies of this review (Appendix 1).

**Summary of Study Characteristics**

Included study characteristics are tabulated in Appendix 2.

**Study design**

Six systematic reviews, two RCTs, four PCSs and twelve RCSs were identified as meeting the selection criteria.

**Population**

The populations in the included studies were patients that underwent various spinal fusion surgeries and patients requiring surgery for nonunion of long bone fractures. Two studies specifically examined the elderly (65 years and older), one RCT examining lumbar fusion surgeries with rhBMP-2 and one RCS examining posterior lumbar fusion surgery (PLF) with rhBMP-2. One RCS examined outcomes with respect to population factors, age, gender, smoking and osteoporosis and other comorbidities.

Specific conditions examined in the included studies that required spinal fusion surgeries were symptomatic single-level degenerative disc disease (DJD), adult ideopathic scoliosis, adult spinal deformity, lumbar spondylosis and lumbar spondylolisthesis. Most studies combined
populations with various spinal pathologies or do not refer to specific pathologies leaving the interventions and outcomes examined consistent.

**Interventions and comparators**

All included studies examined the use of rhBMP-2 and/or rhBMP-7. No studies examined other BMPs for use in spinal surgeries or surgery for long bone trauma. Five of the included systematic reviews examined both rhBMP-2 and rhBMP-7 together, while one systematic review exclusively examined rhBMP-2 in open tibial shaft fractures. Of the four included PCS, one study evaluated rhBMP-7, while the three remaining PCS used rhBMP-2. All of the RCSs included in this review exclusively examined rhBMP-2.

The comparator in the majority of studies was the same surgical procedure as the intervention utilizing autograft instead of BMPs to promote fusion. The comparator in two systematic reviews was autologous iliac crest bone graft (AIBG), while four of these reviews included studies reviewing surgeries with or without bone graft or bone substitutes and without rhBMPs as a comparator. AIBG was used as a control comparator in both of the included RCTs, two of the included PCSs and six of the included RCSs. One PCS, Katayama et al., reported a methodology using a control AIBG for PLF within the same patient, while Lindley et al. used artificial disc replacement (ADR) as a control for ALIF with rhBMP-2. The remaining PCSs and RCSs included studies using other autograft, bone substitutes and/or instrumented surgeries without BMPs as a control.

The spinal surgeries for which BMPs are evaluated for clinical efficacy and/or safety are ALIF, axial lumbar interbody fusion (AxialLIF), PLIF, transforaminal lumbar interbody fusion (TLIF), cervical spinal fusion, adult spinal deformity surgery, extension of previous idiopathic scoliosis fusion to the sacrum, and one report that retrospectively evaluated various and unspecified spinal fusion surgeries.

There were other aspects of the interventions that contained considerable variation in the included studies, including differences in dose, scaffolding material and instrumentation. The dose of BMPs varied not only between studies but sometimes also within treatment groups of the same study. To be effective, BMPs are usually used in conjunction with an osteoconductive matrix to provide scaffolding. Scaffolding material variation existed between studies and sometimes also within studies. The included systematic reviews combined data from studies using different doses, different scaffolding materials and instrumentation in conjunction with BMPs.

Orthopedic surgeries utilizing BMPs for the treatment of long bone fractures and nonunions were examined in one systematic review and one retrospective cohort study. In addition, one systematic review was focused on the use of BMPs in open tibial shaft fractures.

**Outcomes**

The clinical effectiveness outcomes examined in the included studies are successful fusion, fusion rate, operative time, osteogenic score, blood loss, length of hospital stay, reoperation rate, hardware failure, preoperative and postoperative radiographic coronal and sagittal data and Japanese Orthopaedic Association (JOA) score. Patient functional outcomes in five of the included studies are also quantified by Oswestry Disability Index (ODI) and Odom's criteria.
Bone Morphogenetic Proteins

36-Item Short Form Health Survey (SF-36), Scoliosis Research Society (SRS) and health-related quality of life (HRQOL). Safety data for BMPs were the subject of two of the included systematic reviews. The examination of the evidence for an association between BMP use in spinal surgery and an increased risk of cancer was the focus of one of these systematic reviews. Three RCSs hypothesized that retrograde ejaculation is associated with the use of BMPs in ALIF surgery. The complications and adverse events documented in the included studies were cancer, donor-site pain, dural tear, postoperative pain, resorption/osteolysis, extradiscal/ectopic/heterotopic bone formation, graft subsidence, cage migration, dysphagia, dyspnea, neck swelling, respiratory difficulties, antibody responses, infection, hematoma, seroma, radiculitis, inflammatory response, local or systemic toxicity, retrograde ejaculation, postoperative pain, and unspecified total complications.

Two included reports had an economic data component. This component of these reports is not summarized in this review.

Summary of Critical Appraisal

The quality of the included reports varies considerably. A detailed critical appraisal is tabulated in Appendix 4.

Four of the included systematic reviews focus on clinical efficacy and are of moderate quality, limited mostly by the size and number of included studies. All of the included reviews provide clear methodology for a comprehensive literature search however only one systematic review explicitly provides information on a grey literature search. One included systematic review focused on complications and graded the strength of the evidence for its conclusions based on defined criteria, however this systematic review did not assess publication bias or identify studies with potential conflicts of interest (COIs). The focus of another safety-related systematic review focused on incidence of cancer after BMP exposure in spinal surgeries. This review was limited by the data available, could not pool study data due to heterogeneity and was also limited to a short-term cancer risk assessment of two years. Publication bias was assessed in three of the systematic reviews, while COIs were only discussed in one systematic review. The most recent systematic review of clinical efficacy of BMPs in spinal fusion surgeries meeting the selection criteria is from 2009, while the most current systematic review of clinical efficacy for application of BMPs in treatment of long bone trauma was published in 2012.

There were two RCTs that met the selection criteria and are not included in a systematic review. One of these RCTs has a stated COI and is cited in a systematic review examining industry bias in adverse event reporting. This RCT is a multicentre study with clearly defined patient eligibility, clear intervention, outcomes and a follow-up of two years. Both RCTs describe the statistical methods used and both had the assessment of fusion blinded. Neither included RCT included power calculations to determine appropriate sample sizes necessary to detect clinically important effects.

Also included in this review were four nonrandomized prospective controlled trials (PCS). None of the included PCS calculated the statistical power of their results. Significant heterogeneity in the patient groups and unclear, unquantified, inconsistent interventions are present in three of these included PCSs. Stated COIs were present in half of these
studies.\textsuperscript{15,16} One PCS had very consistent interventions and comparators with an average follow-up of five years, however the study evaluated a very small cohort (n=11) of patients.\textsuperscript{14}

Twelve retrospective cohort studies (RCSs) were included in this review.\textsuperscript{17-28} Five of these studies include a stated potential COI and two do not have any statement regarding COI.\textsuperscript{18,19,21-23,27,28} Many of these studies are single centre studies and the interventions were often performed by a single surgeon, which may limit the generalizability of the findings.\textsuperscript{17,18,20,23-26} The quality of some of the included RCSs is limited by inconsistent intervention,\textsuperscript{20,27} comparisons to a noncontemporaneous control,\textsuperscript{18,20,24} and differences between patient treatment groups.\textsuperscript{20-22,27} Strengths of some of the included RCSs include tabulated patient characteristics,\textsuperscript{17,21,23,26} large retrospective cohorts\textsuperscript{20,27,28} and detailed analysis of outcomes.\textsuperscript{21,23,26,27}

One systematic review,\textsuperscript{7} one RCT,\textsuperscript{11} three PCSs\textsuperscript{13,15,16} and five RCSs\textsuperscript{18,19,22,23,27} included a statement that one or more authors had a conflict of interest directly related to one of the interventions being studied. Three additional PCSs and RCSs had no statement regarding COIs.\textsuperscript{14,21,28} Reported adverse events in industry sponsored studies, as compared to independent studies of BMPs is the subject of a systematic review by Carragee et al. published in 2011.\textsuperscript{29} The authors conclude,“Level I and Level II evidence from original FDA summaries, original published data, and subsequent studies suggest possible study design bias in the original trials, as well as a clear increased risk of complications and adverse events to patients receiving rhBMP-2 in spinal fusion. This risk of adverse events associated with rhBMP-2 is 10 to 50 times the original estimates reported in the industry-sponsored peer-reviewed publications.”

Another RCT included in the systematic review by Devine et al. is cited as potentially biased, however, it is the analysis of the data and not the evidence itself that Carragee et al. suggest may contain bias.\textsuperscript{3,29,30} Three other studies cited in the review of Carragee et al. are contained in the included studies or in the meta-analyses of the included systematic reviews, however these remaining citations are all used as examples of a reported increase in adverse events within independent studies and not as biased studies.\textsuperscript{27,29,31,32}

**Summary of Findings**

Key findings of the included studies are summarized in Appendix 4.

Two included systematic reviews examined the literature for evidence of the clinical efficacy of employing BMPs in various spinal fusion surgeries.\textsuperscript{6,9} Both found a statistically significant (P < 0.05) superiority of rhBMP-2 use over autograft in PLF. The more recent systematic review also found a significant superiority of rhBMP-2 use over autograft in ALIF and PLIF.\textsuperscript{6} There was also a significant improvement with rhBMP-2 use over autograft in the rate of fusion.\textsuperscript{9} Despite the evidence favouring rhBMP-2 in fusion, there were no significant differences found in ODI, reoperation rate or clinical failure.\textsuperscript{6,9} Neither systematic review found statistically significant differences between rhBMP-7 and autograft in the spinal fusion surgeries examined.

Evidence for clinical efficacy of employing rhBMP-2 in tibial fractures is examined in two included systematic reviews.\textsuperscript{7,10} Garrison et al. found no significant differences between rhBMP-2 and the use of bone graft substitutes for attaining union, except in acute tibial fracture healing and the rate of hardware failure, both of which statistically favour the use of rhBMP-2.\textsuperscript{7} In open tibial shaft fractures, Wei et al. found evidence for reduced rates of treatment failure and required secondary intervention when rhBMP-2 is used instead of the standard of care. Wei et al. also report no significant differences in adverse events between rhBMP-2 and standard of
Importantly, Garrison et al. highlight the paucity of data for the use of rhBMP-2 in tibial fractures and that the 11 RCTs included in their analysis were evaluated as having a high risk of bias.

One included systematic review specifically quantified cancer occurrence with exposure to BMPs in spinal fusion surgeries. Published in 2012, this systematic review did not find any statistically significant differences in the incidence of cancer between patients exposed to BMPs and those undergoing similar procedures without BMPs. This review did not examine the quality of the included studies or examine existing risk factors within the patient populations. Despite the lack of statistical significance, seven of the small studies included in the review did find an increase cancer incidence in patients exposed to BMPs within a two year follow-up. This data suggests a more detailed assessment, with better analysis of existing risk factors and with sufficient and consistent follow-up is warranted.

Complications associated with BMP use in spinal fusion surgery were examined in one included systematic review. The authors of this review found a high level of evidence identifying commonly reported complications associated with BMP use in lumbar and cervical spine surgery. These include resorption/osteolysis, extradiscal/ectopic/heterotopic bone formation, graft subsidence, graft or cage migration, an elevated antibody response and hematoma. The strength of evidence is graded as moderate with regards to some complication rates in lumbar spine fusion surgeries using BMPs. The complication rates for which there is a moderate strength of evidence are extradiscal/ectopic/heterotopic ossification (3%), graft subsidence (43%), graft migration (2%) in cervical surgery and cage migration (27%), elevated antibody response (rhBMP-2 1%, rhBMP-7 26%) and hematoma (4%) in lumbar surgery. Otherwise, complication rates vary widely, and the evidence for complications in the cervical spine, complication rates in the cervical spine and thoracic spine, as well as dose related complication effects all have low or very low strength of evidence.

Two RCTs included in this review examined the clinical effectiveness of rhBMP-2 use in PLF. Dawson et al. found a statistically significant difference in fusion at 6 months favouring rhBMP-2, however this difference lost significance at 12 and 24 months. Glassman et al. show a statistically significant higher CT grade in PLF using rhBMP-2 over AIBG in patients 65 years and older at a 2 year follow-up. The fusion rate at a two year follow-up, however, was not statistically different between the treatment groups. Other significant findings in these RCTs are shorter operating time and fewer total complications, both favouring rhBMP-2 over AIBG. Both studies showed no other statistically significant differences between rhBMP-2 and AIBG use in PLF, including no significant difference in the functional outcomes, ODI and SF-36, at a two year follow-up. Adverse event rates were reported in Glassman et al. and showed significantly more total complications in the control treatment group although the authors state that none of the complications were directly attributable to either AIBG harvest or the rhBMP-2 use.

Other studies included in this review are nonrandomized prospective controlled studies (PCS) and retrospective cohort studies (RCS). Only accounting for statistically significant findings (P < 0.05) with regards to a fusion related outcome, three of these studies found that rhBMP-2 was superior to autograft, while six studies found no statistically significant differences between rhBMP-2 and autograft in spinal fusion surgeries. None of the included PCSs or RCSs found statistically significant differences favouring autograft in a fusion related outcome. Additionally, none of the included PCSs or RCSs investigated the clinical effectiveness of rhBMP-7 in spinal fusion surgeries.
One RCS examined the use of rhBMP-2 as compared to autograft in the treatment of long bone nonunion. No statistically significant difference was found in surgical success or complication rates, however statistically significant differences included a decrease in operating time and decreased intraoperative blood loss favouring rhBMP-2. Statistically significant differences in adverse events and complications between rhBMP-2 and autograft were reported in the included PCSs and RCSs. Two RCSs reported a statistically significant increase in the incidence of retrograde ejaculation (RE) in male patients undergoing ALIF surgery using BMPs instead of autograft. A subset of the studied cohort was the same in both studies. Another RCS reported no significant difference in the incidence of RE between ALIF with rhBMP-2 and an artificial disc replacement (ADR) procedure using the same surgical approach as ALIF. One PCS and two RCSs reported complications and adverse events specific to rhBMP-2 use in cervical spine fusion surgeries. Both of the large cohort RCSs reported multiple statistically significant differences in the rate of potentially serious complications specific to rhBMP-2 in this context. The complications were wound infection, overall complications, tracheotomies, unplanned intubations after surgery, readmissions, dysphagia, dyspnea and respiratory failure. The smaller cohort PCS found no statistically significant differences in complication rates associated with rhBMP-7 use in cervical spine surgeries. One other PCS and five other RCSs also reported adverse events and complications and found no significant differences between rhBMP-2 and autograft use in spinal fusion surgeries. Most of the references to adverse effects of autograft were qualitative author conclusions however one systematic review included an RCT that reported a donor site pain incidence of 46.5% and one included RCT reported a statistically significant higher incidence of perioperative complications with AIBG.

**Limitations**

There is no consensus in the evidence comparing BMP and autograft use in spinal fusion surgeries. There is also considerable variation in the method in which BMPs are employed in the various surgeries studied in the literature included in this review. There is less evidence identified in this review for applications of rhBMP-7. The literature search strategy employed for this review identified one PCS specifically examining the use of rhBMP-7 in spinal fusion surgery. In addition two systematic reviews examined the evidence for clinical effectiveness of rhBMP-7 in spinal fusion surgery (three RCTs) while one systematic review examined the evidence for rhBMP-7 clinical effectiveness in fracture healing (one RCT).

Of the ten included studies with COIs none reported a statistically significant difference in adverse events between BMPs and autograft except for one study showing a statistically significant increase in adverse events associated anterior cervical fusions using rhBMP-2. Three studies with a statement of no COI found statistically significant increases in adverse events associated with BMP use, two RCSs found an increase in retrograde ejaculation (RE), while one RCS found increases in adverse events associated with BMP use in cervical spine fusion surgeries. Some consideration to the prevalence of conflict of interest within the included literature should be taken into account when reviewing the evidence for any application of BMPs.

Evidence of adverse events reported in this review are limited by the consistency of follow-up and the follow-up time in the included studies. Most notably perhaps is the two year follow-up when examining the incidence of cancers associated with BMP exposure. As the time from...
cancer initiation to clinical detection varies, especially between different types of cancer, it is possible that a two year follow-up is insufficient to fully assess the potential for increased risk.  

CONCLUSIONS AND IMPLICATIONS FOR DECISION OR POLICY MAKING

There is no clear consensus on the evidence of clinical efficacy for the use of rhBMP-2 instead of autograft in spinal fusion surgery. There is no consistency in the dose of rhBMP-2, the instrumentation or scaffolding material to be used in conjunction with rhBMP-2 in the included studies. There is, however, a consensus in the evidence of the retrieved studies of no statistically significant difference in any clinical effectiveness outcome favouring autograft over rhBMP-2, despite the various methods of application. This suggests that the use of rhBMP-2 may be at least as clinically effective as autograft in certain spinal fusion surgeries. Two included systematic reviews and one RCS also demonstrate some advantages of rhBMP-2 use in fusion of long bone fractures.\(^7,10,26\)

No evidence identified in this review supports a statistically significant improvement in a clinical effectiveness outcome for rhBMP-7 over autograft in spinal fusion surgeries or fracture healing. Additionally there were no included studies that directly compared clinical effectiveness of rhBMP-2 with rhBMP-7.

Identified literature contains some evidence for serious acute adverse complications associated with BMP use in cervical spine fusions surgeries.\(^8,27,28\) Evidence from two included RCSs also raise concern of an increased risk of retrograde ejaculation with BMP use in ALIF.\(^17,18\) The increased incidence of cancer associated with BMP use, as summarized by one included systematic review, is not statistically significant.\(^3\) However the studies included in the systematic review are of a small sample size and a relatively short follow-up of two years. There are no statistically significant increases in adverse events or complications for the use of rhBMP-2 in union of long bone fractures,\(^7,10,26\) however evidence for incidence rates of adverse events in this application is low.

PREPARED BY:
Canadian Agency for Drugs and Technologies in Health
Tel: 1-866-898-8439
www.cadth.ca
REFERENCES


APPENDIX 1: SELECTION OF INCLUDED STUDIES

363 citations identified from electronic literature search and screened

308 citations excluded

55 potentially relevant articles retrieved for scrutiny (full text, if available)

56 potentially relevant reports

1 potentially relevant report retrieved from other sources (grey literature, hand search)

32 reports excluded:
- irrelevant comparator (11)
- irrelevant outcomes (2)
- already included in at least one of the selected systematic reviews (12)
- published in language other than English (1)
- other (review articles, retraction, editorials)(6)

24 reports included in review
## APPENDIX 2: SUMMARY OF STUDY CHARACTERISTICS

### Table A2.1: Summary of Study Characteristics of Included Systematic Reviews

<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Study Design, Length of Follow-up</th>
<th>Patient Characteristics, Sample Size (n)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Clinical Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wei et al., 2012(^\text{10})</td>
<td>4 RCTs (mean Jadad score = 3.5) FU=12mos</td>
<td>n=608 (rhBMP=304, control=304) mean age=36.6y</td>
<td>rhBMP-2 in open tibial shaft fractures</td>
<td>Open tibial shaft fractures treated without rhBMP-2</td>
<td>Fusion Time, Reoperation, Infection, Hardware failure, Postoperative pain</td>
</tr>
<tr>
<td>Devine et al., 2012(^\text{3})</td>
<td>7 RCTs 2 FDA safety summaries FU=12-60mos</td>
<td>n=1404 (rhBMP=687, control=717) mean age (yr) range=43-63 mean %male range=45-78</td>
<td>rhBMP-2 or rhBMP-7 in ALIF and PLF</td>
<td>AIBG in ALIF and PLF</td>
<td>Cancer (any type)</td>
</tr>
<tr>
<td>Garrison et al., 2010(^\text{7})</td>
<td>11 RCTs FU=9-12mos</td>
<td>n range=(29-450)</td>
<td>rhBMP-2 or rhBMP-7 with surgery for fracture union</td>
<td>Surgery with or without bone graft or bone substitutes</td>
<td>Fusion, Fusion rate Complications, Hospital Stay, Operative time, Blood loss, Economic data</td>
</tr>
<tr>
<td>Mroz et al., 2010(^\text{8})</td>
<td>7 RCTs 12 PSC 16 RSC FU=2 year max</td>
<td>n=NR mean/complication=1392</td>
<td>rhBMP-2 or rhMBP-7 in ALIF, PLF, PLIF and TLIF</td>
<td>ALIF, PLF, PLIF and TLIF without rhBMP</td>
<td>Complications</td>
</tr>
<tr>
<td>Agarwal et al., 2009(^\text{9})</td>
<td>9 RCTs (median Jadad score =2) 2 PSC 1 RSC FU=12-48mos</td>
<td>n=855 (rhBMP=481, control=374) mean=71 median=44</td>
<td>rhBMP-2 or rhBMP-7 in ALIF, PLF and PLIF</td>
<td>ALIF, PLF and PLIF without rhBMP</td>
<td>Fusion, ODI, Operative time, Blood loss, Hospital stay</td>
</tr>
<tr>
<td>Papakostidis et al., 2008(^\text{9})</td>
<td>6 RCTs 1 PSC mean FU</td>
<td>n=383 mean=48 range= 15-150</td>
<td>rhBMP-2 or rhBMP-7 in PLF</td>
<td>AIBG</td>
<td>Fusion, Fusion rate, Operative time, Reoperation,</td>
</tr>
<tr>
<td>First Author, Publication Year</td>
<td>Study Design, Length of Follow-up</td>
<td>Patient Characteristics, Sample Size (n)</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Clinical Outcomes</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------------------------------</td>
<td>---------------------------------------</td>
<td>--------------</td>
<td>-----------</td>
<td>------------------</td>
</tr>
<tr>
<td></td>
<td>range=12-24mos FU rate range=65-100%</td>
<td>mean age range=42-65yr Comorbidities also tabulated for each included study</td>
<td></td>
<td></td>
<td>Hospital stay</td>
</tr>
</tbody>
</table>

AIBG=autologous iliac crest bone graft; ALIF=anterior lumbar interbody fusion; BMP=bone morphogenetic proteins; FU=follow-up; mos=months ODI=Oswestry Disability Index; PCS=prospective cohort study; PLF=posterolateral lumbar fusion; PLIF=posterior lumbar interbody fusion; RCT=randomized controlled trial; RCS=retrospective cohort study; rhBMP=recombinant human bone morphogenetic protein; TLIF=transforaminal lumbar interbody fusion; yr=years
### Table A2.2: Summary of Study Characteristics of Included Randomized Controlled Trials

<table>
<thead>
<tr>
<th>First Author, Publication Year, Country</th>
<th>Study Design, Length of Follow-up</th>
<th>Patient Characteristics, Sample Size (n)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Clinical Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dawson et al., 2009, USA&lt;sup&gt;11&lt;/sup&gt;</td>
<td>RCT FU=3-24mos (87% at 24mos)</td>
<td>n=40 PPA</td>
<td>rhBMP-2 in single-level PLF w/ instrumentation</td>
<td>AIBG as part of single-level PLF w/ instrumentation</td>
<td>Fusion, Operative time, Blood loss, Hospital stay, ODI, SF-36, Postoperative pain, overall success</td>
</tr>
<tr>
<td>Glassman et al., 2008, USA&lt;sup&gt;12&lt;/sup&gt;</td>
<td>RCT FU=6,12,24mos</td>
<td>n=102 PPA</td>
<td>rhBMP-2 in single or multi-level PLF with instrumentation</td>
<td>AIBG in single or multi-level PLF with instrumentation</td>
<td>Fusion rate, Complications, HRQOL, SF-36, ODI, Economic data, Reoperation, Postoperative pain, Operative time, Blood loss</td>
</tr>
</tbody>
</table>

AIBG = autologous iliac crest bone graft; BMP = bone morphogenetic proteins; FU = follow-up; HRQOL = health-related quality of life; m = months; ODI = Oswestry Disability Index; PCS = prospective cohort study; PLF = posterolateral lumbar fusion; RCT = randomized controlled trial; rhBMP = recombinant human bone morphogenetic protein; SF-36 = 36-Item Short Form Health Survey; yr = years
<table>
<thead>
<tr>
<th>First Author, Publication Year, Country</th>
<th>Study Design, Length of Follow-up</th>
<th>Patient Characteristics, Sample Size (n)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Clinical Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crawford et al., 2010&lt;sup&gt;13&lt;/sup&gt;</td>
<td>Prospective controlled study, FU=2 years</td>
<td>n=60 (rhBMP-2=36 control=24) %Male=7 Mean age, yr rhBMP-2=49.8 control=43.5</td>
<td>rhBMP-2 in extension of previous idiopathic scoliosis fusion to the sacrum</td>
<td>Autograft in extension of previous idiopathic scoliosis fusion to the sacrum</td>
<td>Fusion, Complications, Blood loss, Reoperation, ODI, SRS, Radiographic data (pre and postoperative coronal and sagittal measurements)</td>
</tr>
<tr>
<td>Katayama et al., 2009&lt;sup&gt;14&lt;/sup&gt;</td>
<td>Prospective controlled study, Mean FU=5years</td>
<td>n=11 Internally controlled Posterolateral lumbar fusion %Male=36 Mean age, yr = 56</td>
<td>rhBMP-2 in right side PLF (L4 and L5)</td>
<td>AIBG in left side PLF (L4 and L5)</td>
<td>Fusion, Fusion rate, JOA score, Osteogenic score</td>
</tr>
<tr>
<td>Maeda et al., 2009&lt;sup&gt;16&lt;/sup&gt;</td>
<td>Prospective consecutive controlled cohort, FU=2 year minumum</td>
<td>n=55 (rhBMP-2=23 control=32) %Male=NR Mean age, yr= NR</td>
<td>rhBMP-2 in adult spinal deformity surgery</td>
<td>AIBG in adult spinal deformity surgery</td>
<td>Fusion, Radiographic data (Pre and postoperative coronal and sagittal measurements)</td>
</tr>
<tr>
<td>Leach and Bittar 2009&lt;sup&gt;15&lt;/sup&gt;</td>
<td>Prospective consecutive controlled cohort, a safety assessment, FU=30days</td>
<td>n=131 (rhBMP-7=123 control=8) %Male=NR Mean age, yr=NR</td>
<td>rhBMP-7 in anterior cervical fusion with Ca3(PO4)2 and various instrumentati on</td>
<td>Anterior cervical fusion with Ca3(PO4)2 and various instrumentati on</td>
<td>Complications observable within 30days, Radiographic evidence of prevertebral soft-tissue swelling</td>
</tr>
</tbody>
</table>

AIBG=autologous iliac crest bone graft; DJD=degenerative disc disease; FU=follow-up; HRQOL=health-related quality of life; JOA=Japanese Orthopaedic Association; ODI=Oswestry Disability Index; OR=operating room; PLF=posterolateral lumbar fusion; PLGA=polylactic/glycolic acid; SF-36=36-Item Short Form Health Survey; SRS=Scoliosis Research Society; yr=year
Table A2.4: Summary of Study Characteristics of Included Retrospective Cohort Study

<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Study Design, Length of Follow-up</th>
<th>Patient Characteristics, Sample Size (n)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Clinical Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comer et al., 2012&lt;sup&gt;18&lt;/sup&gt;</td>
<td>Retrospective cohort study FU=2yr</td>
<td>n=472 (rhBMP-2=239 control=233) %Male=100 Mean age, yr rhBMP-2=41.9 control=41.6</td>
<td>rhBMP-2 in ALIF</td>
<td>Local osteophytes or AIBG in ALIF</td>
<td>Complication of retrograde ejaculation</td>
</tr>
<tr>
<td>Hoffmann et al., 2012&lt;sup&gt;20&lt;/sup&gt;</td>
<td>Retrospective cohort study FU=up to 1yr</td>
<td>n=1092 (rhBMP-2=947 control=145) %Male rhBMP-2=40.8 control=51.7 Mean age, yr rhBMP-2=59 control=58</td>
<td>rhBMP-2 in PLF also examined DBM in PLF</td>
<td>autograft in PLF</td>
<td>Fusion, Complications (infection, seroma) Hospital stay, Blood loss</td>
</tr>
<tr>
<td>Lindley et al., 2012&lt;sup&gt;22&lt;/sup&gt;</td>
<td>Retrospective cohort study FU=NR</td>
<td>n=95 (rhBMP-2=54 control=41) %Male=100 Mean age, yr rhBMP-2=49 control=35</td>
<td>rhBMP-2 in ALIF</td>
<td>Artificial disc replacement</td>
<td>Complication of retrograde ejaculation</td>
</tr>
<tr>
<td>Rowan et al., 2012&lt;sup&gt;24&lt;/sup&gt;</td>
<td>Retrospective cohort study FU=3mos</td>
<td>n=104 (rhBMP-2=64 control=40) %Male rhBMP-2=48.4 control=30.0</td>
<td>rhBMP-2 in PLIF</td>
<td>PLIF without rhBMP-2</td>
<td>Immediate and follow-up post-operative leg pain, other complications</td>
</tr>
<tr>
<td>Tessler et al., 2011&lt;sup&gt;26&lt;/sup&gt;</td>
<td>Retrospective cohort study average FU=20.0mos</td>
<td>n=93 (rhBMP-2=19 control=74) %Male rhBMP-2=73.7 control=58.1 Mean age, yr rhBMP-2=41.7 control=45.1</td>
<td>rhBMP-2 in long bone union</td>
<td>AIBG in long bone union</td>
<td>Fusion, Operative time, Blood loss, Hospital stay, Infection</td>
</tr>
<tr>
<td>Gerszten et al., 2011&lt;sup&gt;19&lt;/sup&gt;</td>
<td>Retrospective cohort study FU=2+yr</td>
<td>n=99 (rhBMP-2=45 control=56) %Male=44</td>
<td>rhBMP-2 in AxiaLIF</td>
<td>AIBG in AxiaLIF</td>
<td>Fusion, Functional outcomes rhBMP-2 by ODI</td>
</tr>
<tr>
<td>First Author, Publication Year</td>
<td>Study Design, Length of Follow-up</td>
<td>Patient Characteristics, Sample Size (n)</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Clinical Outcomes</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>----------------------------------</td>
<td>----------------------------------------</td>
<td>--------------</td>
<td>------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Carragee et al., 2011&lt;sup&gt;17&lt;/sup&gt;</td>
<td>Retrospective cohort study, FU=1yr</td>
<td>n=243 (rhBMP-2=69 control=174) %Male=100 Mean age, yr rhBMP-2=42.4 control=40.9</td>
<td>rhBMP-2 in ALIF Local osteophytes or AIBG in ALIF</td>
<td>control by Odom’s criteria</td>
<td>Complication of retrograde ejaculation</td>
</tr>
<tr>
<td>Williams et al., 2011&lt;sup&gt;27&lt;/sup&gt;</td>
<td>Retrospective cohort study, FU=N R</td>
<td>n=55862 (rhBMP-2=11933 control=43929) %Male=N R Mean age, yr rhBMP-2=52.1 control=36.4</td>
<td>rhBMP-2 in any spinal fusion procedure Spinal fusion without rhBMP-2</td>
<td>Complications (infections, hematomas, seromas)</td>
<td></td>
</tr>
<tr>
<td>Lee et al., 2010&lt;sup&gt;21&lt;/sup&gt;</td>
<td>Retrospective cohort study, FU=2yr</td>
<td>n=75 (rhBMP-2, &gt;65yr=34, &lt;65yr=52 control=41) Mean age, yr &gt;65yr=74.1 &lt;65yr=49.9 control=72.4</td>
<td>rhBMP-2 in PLF in &gt;65yr and rhBMP-2 in PLF in &lt;65yr autograft in PLF in &gt;65yr</td>
<td>Fusion, Fusion rate, Outcomes with regards to age, gender, nicotine use, osteoporosis Complications</td>
<td></td>
</tr>
<tr>
<td>Yaremchuk et al., 2010&lt;sup&gt;28&lt;/sup&gt;</td>
<td>Retrospective cohort study, FU=30 days</td>
<td>n=775 (rhBMP-2=260 control=515) %Male=N R Mean age=N R</td>
<td>rhBMP-2 in cervical spine fusion Cervical spine fusion without rhBMP-2</td>
<td>Complications</td>
<td>Airway complications, Hospital stay, Economic data and death within 90days</td>
</tr>
<tr>
<td>Taghavi et al., 2010&lt;sup&gt;25&lt;/sup&gt;</td>
<td>Retrospective cohort study, FU&gt;2yr</td>
<td>n=44 (rhBMP-2=24 control=20) %Male rhBMP-2=45.8 control=55.0 Mean age, yr rhBMP-2=57.3 control=55.8</td>
<td>rhBMP-2 in revision PLF also examined BMAA in revision PLF AIBG in revision PLF</td>
<td>Fusion, Fusion Rate, Complications</td>
<td></td>
</tr>
<tr>
<td>First Author, Publication Year</td>
<td>Study Design, Length of Follow-up</td>
<td>Patient Characteristics, Sample Size (n)</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Clinical Outcomes</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------------------</td>
<td>------------------------------------------</td>
<td>--------------</td>
<td>------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Rihn et al., 2009&lt;sup&gt;23&lt;/sup&gt;</td>
<td>Retrospective cohort study average FU=19.1 mos</td>
<td>n=119 (rhBMP-2=86 control=33) %Male=52.9 Mean age =47.4 yr</td>
<td>rhBMP-2 in single-level TLIF</td>
<td>AIBG in single-level TLIF</td>
<td>Fusion, Complications</td>
</tr>
</tbody>
</table>

**AIBG**=autologous iliac crest bone graft; **ALIF**=anterior lumbar interbody fusion; **AxiaLIF**=axial lumbar interbody fusion; **BMAA**=bone marrow aspirate with allograft; **BMP**=bone morphogenetic proteins; **DBM**=demineralized bone matrix; **FU**=follow-up; **mos**=months; **NR**=not reported; **ODI**=Oswestry Disability Index; **PLF**=posterior lumbar fusion; **rhBMP**=recombinant human bone morphogenetic protein; **TLIF**=transforaminal lumbar interbody fusion; **yr**=years
### APPENDIX 3: SUMMARY OF CRITICAL APPRAISAL

<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Systematic Reviews (6)</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Wei et al., 2012<sup>10</sup>  | ● Statement of no relevant COI  
● Literature search selection/exclusion methodology clearly outlined  
● Study quality assessed (Jadad scale)  
● Data extraction performed  
● Tests of statistical heterogeneity were conducted and a random effects model was used when heterogeneity was evident  
● Table of included studies  
● Examined multiple relevant endpoints separately  
● Quantified conclusions | ● Possible short-term bias  
● Publication bias not assessed |
| Devine et al., 2012<sup>3</sup> | ● Systematically reviews adverse cancer events  
● Literature search, selection/exclusion methodology clearly outlined  
● Discussion of COIs in studies  
● Statement of no COI                                                                 | ● Evaluated relatively short-term risk of cancer (2 years)  
● Study quality not objectively assessed  
● Data not extracted  
● Patient characteristics in included studies not examined – no mention of risk factors in patient groups  
● Very little data available |
| Garrison et al., 2010<sup>7</sup> | ● Literature search methodology clearly described  
● Publication risk of bias assessed (Cochrane Collaboration’s ‘Risk of bias’ assessment tool)  
● Data extraction performed  
● Methodological quality assessed and tabulated  
● Tests of statistical heterogeneity were conducted  
● Quantified conclusions  
● Tables of included and excluded studies (reason for exclusion tabulated) | ● No predefined exclusion criteria  
● Stated COI |
| Mroz et al., 2010<sup>8</sup>   | ● Literature search methodology described  
● Inclusion/exclusion criteria tabulated  
● Data extraction performed  
● Level of evidence on pooled data graded using defined criteria  
● Data on complications pooled and tabulated  
● Statement of no COI                                                                 | ● Publication bias not assessed  
● Studies with COIs not identified  
● No tests of statistical heterogeneity  
● Patient characteristics NR |
| Agarwal et al., 2009<sup>6</sup> | ● Literature search inclusion/exclusion methodology clearly outlined  
● Data extraction methodology outlined                                                                 | ● Study from 2009  
● Patient characteristics in included trials not examined |
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
</table>
| Papakostidis et al., 2008⁹ | ● Literature search selection/exclusion methodology clearly outlined  
● Scientific methodology quantitatively assessed (PEDro scale)  
● Tests of statistical heterogeneity were conducted  
● Patient characteristics in included studies examined  
● Publication bias assessed (funnel plots)  
● Quantified conclusions  
● Examined multiple relevant outcomes separately  
● Statement of no COI | ● Study from 2008  
● Study quality not reflected in strength of conclusions  
● Data combined in the presence of statistical heterogeneity |
| Randomized Controlled Trials (2) | | |
| Dawson et al., 2009¹¹ | ● Multicenter study  
● Clearly defined patient eligibility, intervention and outcomes  
● Statistical methods described  
● Patient group characteristics tabulated  
● Radiographic assessment blinded  
● Follow-up 24 months | ● No statistical power calculation based upon sample size  
● Cited as evidence for possible industry bias²⁹  
● Stated COI |
| Glassman et al., 2008¹² | ● Patient group characteristics tabulated (60y+)  
● Outcomes including HRQOL and economic measures defined  
● Statement of no COI  
● Fusion assessment blinded  
● Statistical methods described | ● Conducted at a single centre  
● Discretionary use of bone graft extenders in both treatment groups  
● No statistical power calculation based upon sample size |
| Nonrandomized Prospective Controlled Studies (4) | | |
| Crawford et al., 2010¹³ | ● Follow-up > 3 years  
● Detailed pre and postoperative radiographic data  
● Special patient population (previous idiopathic scoliosis fusion to the | ● Single centre study  
● Significant heterogeneous patient groups with respect to age, number of anterior levels fused, number of thoracoabdominal approaches and |
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
</table>
| Katayama et al., 2009<sup>14</sup> | - Intervention and control in the same patients  
- Follow-up average 5 years  
- Detailed fusion scores tabulated  
- Consistent intervention and comparator | - Small cohort  
- Single centre study  
- No statistical power calculation based upon sample size  
- No COI statement |
| Maeda et al., 2009<sup>16</sup> | - Patient group characteristics tabulated  
- Statistical methods described  
- Radiographic outcomes tabulated | - Statistically significant difference in patient groups follow-up (2.7 vs 4.9 years), number of fused vertebrae, and surgical approach  
- Inconsistent rhBMP-2 dose, surgical approach, use of segmental pedicle screws and local bone graft harvesting with BMP  
- No statistical power calculation based upon sample size  
- Stated COI |
| Leach and Bittar 2009<sup>15</sup> | - Radiographic assessment blinded | - Interventions conducted by a single surgeon  
- Stated COI  
- Patient groups not balanced  
- Patient group characteristics NR  
- Retrospective control group for prevertebral soft-tissue swelling  
- Unclear and unquantified interventions (dose, instrumentation)  
- No statistical power calculation based upon sample size  
- Unquantified complication outcomes |
| Retrospective Cohort Studies (12) | | |
| Comer et al., 2012<sup>18</sup> | - Patient group characteristics balanced and tabulated  
- Research question, intervention and outcome well defined  
- Statistical methods described  
- Blinded relevant diagnosis  
- Thorough follow-up | - Single centre study  
- Stated COI  
- Complications not compared to contemporaneous control |
| Hoffmann et al., 2012<sup>20</sup> | - Large cohort  
- Patient group characteristics tabulated  
- Statistical methods described  
- Statement of no COI | - Single centre study  
- Statistically significant differences between groups in some patient characteristics  
- Inconsistent intervention  
- Surgical outcome and complications not compared to contemporaneous control |
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
</table>
| Lindley et al., 2012<sup>22</sup> | ● Patient group characteristics tabulated  
● Research question intervention and outcome defined  
● Statistical methods described | ● Complications lack definitive diagnosis retrospectively  
● Statistically significant differences between patient groups characteristics and surgical procedures  
● Retrospective study likely underestimates the specific outcome<sup>18</sup>  
● Stated COI  
● No reference to length of follow-up  
● Complication outcome diagnosed by questionnaire |
| Rowan et al., 2012<sup>24</sup> | ● Patient group characteristics tabulated and matched (except for previous surgery)  
● Research question and outcomes well defined  
● Statistical methods described  
● Statement of no COI | ● Single centre and surgeon  
● Surgical outcome and complications not compared to contemporaneous control  
● Results do not strongly support conclusions |
| Tressler et al., 2011<sup>26</sup> | ● Patient inclusion and exclusion criteria tabulated  
● Patient nonunion characteristics tabulated and matched  
● Patient group characteristics matched  
● Research question and outcomes clearly defined  
● Statistical methods described  
● Statement of no COI  
● Multiple surgeons using consistent intervention  
● Outcomes examined with respect to age and other risk factors  
● Power analysis performed (40%) | ● BMP used in conjunction with allograft  
● Large treatment group size difference |
| Gerszten et al., 2011<sup>19</sup> | ● Patient group characteristics tabulated  
● Research question and intervention defined  
● Statistical methods described | ● Stated COI  
● Outcome quantification method is different between treatment groups  
● Results do not strongly support conclusions |
| Carragee et al., 2011<sup>17</sup> | ● Patient group characteristics tabulated  
● Research question and outcomes defined  
● Statement of no COI  
● Statistical methods described | ● Unclear how complication outcome was diagnosed  
● Single centre and surgeon |
| Williams et al., 2011<sup>27</sup> | ● Large retrospective cohort  
● Patient group characteristics tabulated  
● Research question and outcomes well defined  
● Outcomes for different surgical procedures and diagnosis examined | ● Statistically significant differences between patient groups characteristics  
● No defined follow-up timeframe  
● Stated COI  
● Inconsistent intervention |
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>independently</td>
<td></td>
<td></td>
</tr>
<tr>
<td>● Statistical methods described</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee et al., 2010&lt;sup&gt;21&lt;/sup&gt;</td>
<td>● Patient group characteristics tabulated</td>
<td>● Differences between patient groups (significance not calculated)</td>
</tr>
<tr>
<td></td>
<td>● Research question intervention and outcome defined</td>
<td>● No COI statement</td>
</tr>
<tr>
<td></td>
<td>● Statistical methods described</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Follow-up &gt; 3 years</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Outcomes examined with respect to age and other risk factors</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Radiographic assessment blinded</td>
<td></td>
</tr>
<tr>
<td>Yaremchuk et al., 2010&lt;sup&gt;28&lt;/sup&gt;</td>
<td>● Large retrospective cohort</td>
<td>● No COI statement</td>
</tr>
<tr>
<td></td>
<td>● Statistical methods described</td>
<td>● Patient characteristics NR</td>
</tr>
<tr>
<td></td>
<td>● Specific research question</td>
<td>● Methodology of cohort matching NR</td>
</tr>
<tr>
<td></td>
<td>● Outcomes well defined</td>
<td>● Acute outcomes only, short follow-up</td>
</tr>
<tr>
<td></td>
<td>● High resolution follow-up for some outcomes</td>
<td></td>
</tr>
<tr>
<td>Taghavi et al., 2010&lt;sup&gt;25&lt;/sup&gt;</td>
<td>● Patient group characteristics balanced and tabulated</td>
<td>● Single centre and surgeon</td>
</tr>
<tr>
<td></td>
<td>● Statement of no COI</td>
<td>● Low number of nonunions limits predictive power of study</td>
</tr>
<tr>
<td></td>
<td>● Research question intervention and outcome defined</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Statistical methods described</td>
<td></td>
</tr>
<tr>
<td>Rihn et al., 2009&lt;sup&gt;23&lt;/sup&gt;</td>
<td>● Statistical methods described</td>
<td>● Single centre study</td>
</tr>
<tr>
<td></td>
<td>● Consistent surgical method</td>
<td>● Stated COI</td>
</tr>
<tr>
<td></td>
<td>● Research question defined</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● Complication data tabulated</td>
<td></td>
</tr>
</tbody>
</table>

BMP = bone morphogenetic proteins; COI = Conflict of Interest; HA = hydroxyapatite; HRQOL = health-related quality of life; NR = Not reported
APPENDIX 4: SUMMARY OF FINDINGS

<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Main Study Findings</th>
<th>Author's Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Systematic Reviews (6)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wei et al., 2012(^{10})</td>
<td><strong>Open tibial shaft fractures</strong></td>
<td>Clinical Effectiveness</td>
</tr>
<tr>
<td></td>
<td>RR&lt;1 favours BMP</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Secondary Intervention</strong></td>
<td>Clinical Effectiveness</td>
</tr>
<tr>
<td></td>
<td>rhBMP-2 vs standard of care (2 RCTs)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR (95%CI): 0.65 (0.48, 0.89) (I(^2)=44%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Treatment Failure</strong></td>
<td>Clinical Effectiveness</td>
</tr>
<tr>
<td></td>
<td>rhBMP-2 vs standard of care (2 RCTs)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR (95%CI): 0.62 (0.48, 0.81) (I(^2)=0%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Hardware Failure</strong></td>
<td>Clinical Effectiveness</td>
</tr>
<tr>
<td></td>
<td>rhBMP-2 vs standard of care (2 RCTs)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR (95%CI): 0.76 (0.35, 1.68) (I(^2)=76%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>RR&lt;1 favours Standard of Care</strong></td>
<td>Clinical Effectiveness</td>
</tr>
<tr>
<td></td>
<td>Fracture Healing Rate in 20 weeks</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rhBMP-2 vs standard of care (2 RCTs)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR (95%CI): 1.35 (0.74, 2.47) (I(^2)=92%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Adverse Events</strong></td>
<td>Clinical Effectiveness</td>
</tr>
<tr>
<td></td>
<td>RR&lt;1 favours BMP</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Infection</strong></td>
<td>Clinical Effectiveness</td>
</tr>
<tr>
<td></td>
<td>rhBMP-2 vs standard of care (2 RCTs)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR (95%CI): 1.19 (0.57, 2.49) (I(^2)=76%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Pain</strong></td>
<td>Clinical Effectiveness</td>
</tr>
<tr>
<td></td>
<td>rhBMP-2 vs standard of care (2 RCTs)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RR (95%CI): 0.96 (0.75, 1.22) (I(^2)=76%)</td>
<td></td>
</tr>
<tr>
<td>Devine et al., 2012(^{3})</td>
<td><strong>Adverse Events -Cancer only</strong></td>
<td>Clinical Effectiveness</td>
</tr>
<tr>
<td></td>
<td>rhBMP-7</td>
<td>Not examined</td>
</tr>
<tr>
<td></td>
<td>3 pilot studies (n=36)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>12.5% cancer in rhBMP-7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.3% in controls</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 RCT (n=34)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.6%) 1 case of cancer in rhBMP-7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0% in controls</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rhBMP-2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 RCT (n=463)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Clinical Effectiveness</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Not examined</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Adverse Events -Cancer only</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

"The findings of the current study confirm that rhBMP-2 in an absorbable collagen sponge could decrease the secondary intervention rate and the treatment failure rate when used with intramedullary nail fixation for the treatment of open tibial shaft fractures. "(pp. 853)

"...no significant difference was seen in infection rate between groups." (pp. 851)
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Main Study Findings</th>
<th>Author’s Conclusions</th>
</tr>
</thead>
</table>
| Garrison et al., 2010<sup>7</sup> | **Clinical Effectiveness**<br>**Tibial fracture attaining union without secondary procedure**<br>rhBMP-2 vs Bone graft substitutes (4 RCTs)<br>RR (95%CI): 1.19 (0.99, 1.43) (I²=32%)<br>RR>1 favours rhBMP-2<br><br>**Prior nonunion of the long bones attaining union**<br>rhBMP-2 vs Bone graft substitutes (6 RCTs)<br>RR (95%CI): 1.02 (0.93, 1.48) (I²=39%)<br>RR>1 favours rhBMP-2<br><br>**Union of distal radial fractures with symptomatic malunion**<br>rhBMP-7 vs Bone graft substitutes (1 RCT)<br>RR (95%CI): 0.76 (0.53, 1.09)<br>RR>1 favours rhBMP-7<br><br>**Acute fracture: requirement for revision surgery to attain union**<br>rhBMP-2 vs Bone graft substitutes (4 RCTs)<br>RR (95%CI): 0.65 (0.50, 0.83) (I²=6%)<br>RR<1 favours rhBMP-2<br><br>**Nonunion of tibia or other long bone requiring revision surgery to attain union**<br>rhBMP-2 vs Bone graft substitutes (2 RCTs)<br>RR (95%CI): 0.41 (0.13, 1.28) (I²=0%)<br>RR<1 favours rhBMP-2 | 35)  
“Cancer risk with BMP-2 may be dose dependent, illustrating the need to continue to study this technology and obtain longer follow-up on patients currently enrolled in the FDA trials.” (pp. 35)  

Clinical Effectiveness  
“This review highlights a paucity of data on the use of BMP in fracture healing as well as considerable industry involvement in currently available evidence. There is limited evidence to suggest that BMP may be more effective than controls for acute tibial fracture healing, however, the use of BMP for treating nonunion remains unclear.” (pp. 23)  

Adverse Events  
“Trial participants who received BMP experienced similar adverse effects to those no receiving BMP (infection, hardware failure, heterotopic bone formation and immunogenic reactions). However, patients given BMP instead of bone autografts will have avoided problems associated with extraction of the bone from another site in their body.” (pp. 2)
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Main Study Findings</th>
<th>Author's Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acute fracture: Hardware failure</strong>&lt;br&gt;rhBMP-2 vs Bone graft substitutes (3 RCTs)&lt;br&gt;RR (95%CI): 0.64 (0.42, 0.96) (I²=0%)&lt;br&gt;RR&lt;1 favours rhBMP-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mroz et al., 2010⁸</td>
<td><strong>Adverse Events</strong>&lt;br&gt;<strong>Strength of Evidence: High</strong>&lt;br&gt;Common complications in lumbar and cervical spine fusion surgery with BMP:&lt;br&gt;resorption/osteolysis, extradiscal/ectopic/heterotopic bone formation, graft subsidence, graft or cage migration, antibody response and hematoma</td>
<td><strong>Clinical Effectiveness</strong>&lt;br&gt;Not examined</td>
</tr>
<tr>
<td></td>
<td><strong>Strength of Evidence: Moderate</strong>&lt;br&gt;Complication rates in lumbar and cervical spine fusion surgery with BMP:&lt;br&gt;Events in narrow range:&lt;br&gt;cervical surgery: extradiscal/ectopic/heterotopic ossification (3%), graft subsidence (43%), graft migration (2%), lumbar surgery: cage migration (27%), antibody response (rhBMP-2 1%, rhBMP-7 26%) and hematoma (4%)&lt;br&gt;Otherwise rates vary widely among studies</td>
<td><strong>Adverse Events</strong>&lt;br&gt;“Clinical Recommendations”&lt;br&gt;1. Given the potential complications related to the use of BMP-2 in ventral cervical spine surgery, its use is not recommended until its clinical efficacy and safety is adequately defined by well-designed and executed studies.</td>
</tr>
<tr>
<td></td>
<td><strong>Strength of Evidence: Low</strong>&lt;br&gt;Common complications unique to cervical spine fusion surgery with BMP:&lt;br&gt;dysphagia, neck swelling and respiratory difficulties&lt;br&gt;Complication rates in cervical spine fusion surgery</td>
<td>2. Given the potential complications related to the use of BMP-2 in posterior lumbar interbody fusion, there are concerns with regard to its routine use in this fusion application.</td>
</tr>
<tr>
<td></td>
<td><strong>Strength of Evidence: Very Low</strong>&lt;br&gt;Common complications in thoracic spine fusion surgery with BMP:&lt;br&gt;dysphagia, neck swelling, respiratory difficulties, and wound complications&lt;br&gt;Complication rates in thoracic spine fusion surgery&lt;br&gt;0.8% wound complications, neck swelling and dysphagia&lt;br&gt;4.7% hematomas, seromas, infections or dehiscence&lt;br&gt;Dose relationship associated with BMP complications</td>
<td>3. There is insufficient data to validate the use of BMP-2 for posterior cervical or thoracic fusion surgery.” (pp. 102)</td>
</tr>
<tr>
<td>First Author, Publication Year</td>
<td>Main Study Findings</td>
<td>Author’s Conclusions</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------</td>
<td>----------------------</td>
</tr>
</tbody>
</table>
| **Agarwal et al., 2009**<sup>6</sup> | **Clinical Effectiveness**  
RR<1 favours BMP  
*rhBMP-2*  
**Radiographic nonunion at 12-24 months**  
rhBMP-2 vs AIBG for ALIF (3 RCTs)  
RR (95%CI): 0.28 (0.13, 0.57) (I<sup>2</sup>=51.4%)  
rhBMP-2 vs AIBG for PLF/PLIF (3RCTs)  
RR (95%CI): 0.27 (0.13, 0.57) (I<sup>2</sup>=2.0%)  
**ODI at 12-24 months**  
rhBMP-2 vs AIBG for ALIF (2 RCTs)  
RR (95%CI): 0.84 (0.48, 1.48) (I<sup>2</sup>=0%)  
rhBMP-2 vs AIBG for PLF/PLIF (2 RCTs)  
RR (95%CI): 0.74 (0.40, 1.35) (I<sup>2</sup>=0%)  | **Clinical Effectiveness**  
"Recombinant human BMP-2 may be an effective alternative to AIGB in lumbar fusion. Data are limited for other bone graft substitutes." (pp. 729)  
"Our review suggests that rhBMP-2 is an effective tool to facilitate lumbar fusion in single-level lumbar DJD. Given the extent of improvement in radiographic nonunion and the trend toward a favorable benefit for clinical outcomes, it may be considered an effective alternative to autografts and allografts in lumbar fusion." (pp. 739) |
| **Papakostidis et al., 2008**<sup>9</sup> | **Clinical Effectiveness**  
RR<1 favours BMP  
*rhBMP-7*  
**Radiographic nonunion at 12-48 months**  
rhBMP-7 vs autograft for PLF (3 RCTs)  
RR (95%CI): 1.09 (0.44, 2.73) (I<sup>2</sup>=0%)  
**Adverse Events**  
No adverse events specific to intervention summarized. One RCT reported 46.5% donor site pain for AIBG controls.  | **Clinical Effectiveness**  
"Although the radiographic results appeared better in the group of BMPs, the exact role of type, dose and carrier of BMPs and the cost-effectiveness of their use need further clinical delineation." (pp. 680)  
*Use of BMPs in the setting of instrumented
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Main Study Findings</th>
<th>Author’s Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR (95%CI): 0.45 (0.31, 0.64) ($I^2=58.3%$)</td>
<td>posterolateral fusion produces superior radiographic results of solid fusion over noninstrumented procedures.</td>
</tr>
</tbody>
</table>
|                               | *Failure at 12 months*  
BMPs vs AIBG for PLF (5 RCTs, 1PCS)  
RR (95%CI): 0.50 (0.30, 0.82) ($I^2=55.0\%$) | ● rhBMP-2 seems to be more efficient over autologous iliac crest bone graft with respect to successful fusion.  
● OP-1 seems to be equivalent to autologous ICBG with respect to fusion success.  
● Use of BMPs is associated with shortened hospitalization of patients undergoing posterolateral spinal fusion compared with the use of autologous ICBG.” (pp. 691) |
|                               | *Failure at 24-36 months*  
BMPs vs AIBG for PLF (4 RCTs, 1PCS)  
RR (95%CI): 0.42 (0.26, 0.68) ($I^2=12.9\%$) |  |
|                               | *Reoperation*  
rhBMP-2 vs autograft for PLF (3 RCTs, 1PCS)  
RR (95%CI): 0.46 (0.14, 1.49) ($I^2=0\%$) |  |
|                               | rhBMP-7 vs autograft for PLF (1RCTs)  
RR (95%CI): 2.00 (0.21, 18.69) |  |
|                               | *Clinical Failure*  
rhBMP-2 vs autograft for PLF (2 RCTs)  
RR (95%CI): 0.86 (0.29, 2.57) ($I^2=4.7\%$) |  |
| **Randomized Controlled Trials (2)** |  |  |
| Dawson et al., 2009$^{11}$ | **Clinical Effectiveness**  
Fusion for PLF  
6 months ($P = 0.032$)  
rhBMP-2 (95%CI): 91%(71%, 99%)  
AIBG (95%CI): 58%(33%, 80%) | **Clinical Effectiveness**  
“Compared with an iliac crest bone graft, the combination of an absorbable collagen sponge soaked with rhBMP-2 and ceramic granules resulted in trends toward improvements in clinical outcomes and toward a higher rate of radiographic fusion. This combination of an osteoinductive agent with an osteoconductive matrix may be an effective replacement for autograft in single-level posterolateral lumbar fusion.” |
|                               | 12 months ($P = 0.184$)  
rhBMP-2 (95%CI): 89%(67%, 99%)  
AIBG (95%CI): 65%(41%, 85%) |  |
|                               | 24 months ($P = 0.120$)  
rhBMP-2 (95%CI): 95%(75%, 100%)  
AIBG (95%CI): 67%(41%, 87%) |  |
|                               | **Surgical Data**  
Operative time (hrs) ($P = 0.415$)  
rhBMP-2 avg ± std dev (95%CI): 2.4 ± 0.7 (2.1, 2.7)  
AIBG avg ± std dev(95%CI): |  |
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Main Study Findings</th>
<th>Author’s Conclusions</th>
</tr>
</thead>
</table>
| **Glassman et al., 2008**¹² | **Clinical Effectiveness**  
PLF (age > 65 years)  
2 year postoperative CT grade (P = 0.030)  
rhBMP-2 = 4.3 ± 1.3  
AIBG = 3.8 ± 0.9  

Operative time (min) (P = 0.024)  
rhBMP-2 = 248 ± 58.8  
AIBG = 270 ± 33.6  

Statistically significant improvements in both groups. No statistical difference between groups including the following:  
2 year mean change in ODI score (P > 0.05)  
rhBMP-2 = 15.8 ± 17.7  
AIBG = 13.0 ± 15.5  

2 year mean change in SF-36 score (P > 0.05)  
rhBMP-2 = 6.6 ± 9.3  
AIBG = 7.5 ± 8.4  

2 year fusion rate (P > 0.05)  
rhBMP-2 = 86.3%  
AIBG = 70.8%  

**Adverse Events**  
Lumbar Spine Fusion (age > 65 years)  
Total complications (n/N) (P = 0.014)  
rhBMP-2 = 8/50 | Clinical Effectiveness  
“RhBMP-2/ACS is a viable AIBG replacement in older patients in terms of safety, clinical efficacy, and cost-effectiveness.” (pp. 2843)  

**Adverse Events**  
“Adverse effects with rhBMP-2 specific to older patients have never been reported, and no indication of any rhBMP-2-related complications was detected in this study. In fact, a significantly higher incidence of perioperative complications was observed in the AIBG group.” (pp. 2848) |
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Main Study Findings</th>
<th>Author's Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AIBG = 20/52</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NRS leg pain (P = 0.031)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rhBMP-2 = 8.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AIBG = 7.2</td>
<td></td>
</tr>
</tbody>
</table>

**Nonrandomized Prospective Controlled Studies (4)**

<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Clinical Effectiveness</th>
<th>Adverse Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crawford et al., 2010</td>
<td><strong>Clinical Effectiveness</strong>&lt;br&gt;Extension of an existing idiopathic scoliosis fusion to the sacrum&lt;br&gt;Overall there were no clinically and statistically significant differences found between rhBMP-2 and autograft treatment groups. Specifically however:&lt;br&gt;Posterior fusion grade from L4 to the sacrum was better in rhBMP-2 group (2.3 ± 0.7 vs 1.7 ± 0.9; P = 0.021)&lt;br&gt;<strong>Adverse Events</strong>&lt;br&gt;Overall there were no clinically and statistically significant differences found between rhBMP-2 and autograft treatment goups. Follow-up complications favouring rhBMP-2 approached statistical significance (rhBMP-2= 9/36 vs autograft= 12/24; P = 0.058)</td>
<td><strong>Clinical Effectiveness</strong>&lt;br&gt;“BMP-2 is a safe and effective alternative to iliac or rib harvesting when extending an existing idiopathic scoliosis fusion to the sacrum” (pp. 1843)&lt;br&gt;<strong>Adverse Events</strong>&lt;br&gt;“We did not observe an increase in the adverse events that have been associated with rhBMP-2 use.” (pp. 1848)</td>
</tr>
<tr>
<td>Katayama et al., 2009</td>
<td><strong>Clinical Effectiveness</strong>&lt;br&gt;PLF&lt;br&gt;The osteogenic score favoured autograft at 6 months postoperation with statistical significance. At 12 and 24 months after surgery this difference was not apparent. Fusion progressed until 2 years after surgery while no change was observed after that time.&lt;br&gt;<strong>Adverse Events</strong>&lt;br&gt;Methodology limited the examination of adverse events specific to treatment.</td>
<td><strong>Clinical Effectiveness</strong>&lt;br&gt;“There were no statistically significant differences between rhBMP-2 and autogenous bone grafts. rhBMP-2 can be used as the sole source of osteogenesis with success equivalent to an autologous graft of the PLF.” (pp. 1066)&lt;br&gt;<strong>Adverse Events</strong>&lt;br&gt;“The use of rhBMP-2 would eliminate the harvesting of autologous bone graft from a patient and the resulting potential complications.” (pp. 1066)</td>
</tr>
<tr>
<td>First Author, Publication Year</td>
<td>Main Study Findings</td>
<td>Author's Conclusions</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------</td>
<td>-----------------------</td>
</tr>
</tbody>
</table>
| **Maeda et al., 2009** | **Clinical Effectiveness**  
Adult spinal deformity surgery  
*Fusion (P = 0.057)*  
rhBMP-2 (n/N) = 95.7% (22/23)  
AIBG (n/N) = 71.9% (23/32)  
Average radiographically measured correction rate of coronal and sagittal Cobb angles favoured rhBMP-2 over autograft (50.6% vs 42.5%; P = 0.077) | **Clinical Effectiveness**  
"The pseudarthrosis rate observed in the current study in the BMP group (4.3%) compares favorably to pseudarthrosis rate in the AIBG group (28.1%)."  
(pp. 2212)  
**Adverse Events**  
No events attributable to rhBMP-2. |
| **Leach and Bittar 2009** | **Clinical Effectiveness**  
Not examined in this study | **Clinical Effectiveness**  
"The effect of BMP-7 on the rate and timing of fusion, as well as clinical outcome, is yet to be elucidated."  
(pp. 1417)  
**Adverse Events**  
"We concluded that BMP-7 can be used safely in anterior cervical fusion."  
(pp. 1417) |
| **Comer et al., 2012** | **Clinical Effectiveness**  
Not examined in this study | **Clinical Effectiveness**  
Not examined in this study.  
**Adverse Events**  
Retrograde Ejaculation in ALIF (P = 0.0012)  
rhBMP-2 (90%CI): 6.3%(3.71%, 8.89%)  
Autograft (90%CI): 0.9%(-0.14%, 1.85%) | **Adverse Events**  
"This study confirms previous reports of a higher rate of RE in ALIF procedures using rhBMP-2 and an open anterior approach to the spine."  
(pp. 882) |
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Main Study Findings</th>
<th>Author's Conclusions</th>
</tr>
</thead>
</table>
| Hoffmann et al., 2012<sup>20</sup> | **Clinical Effectiveness**  
PLF  
Nonunions ($P < 0.001$)  
rhBMP-2 (n/N) = 4.3% (41/947)  
Autograft (n/N) = 15.2% (22/145)  

$Hospital Stay (days)$ ($P = 0.006$)  
rhBMP-2 mean (range) = 4.9 (0-24)  
Autograft mean (range) = 5.8 (0-47)  

Overall there were no other clinically and statistically significant differences in outcomes found between rhBMP-2 and autograft treatment groups.  

**Adverse Events**  
Overall there were no clinically and statistically significant differences in adverse events found between rhBMP-2 and autograft treatment groups. |  
**Clinical Effectiveness**  
"rhBMP-2 supplementation instead of AIBG or bone marrow aspirate results in higher fusion rates compared to autograft alone or autograft plus DBM." (pp. 1105)  

**Adverse Events**  
"we have seen a slight increase in seroma formation for rhBMP-2 procedures, but this was not statistically significant." (pp.1109) |
| Lindley et al., 2012<sup>22</sup> | **Clinical Effectiveness**  
Not examined in this study |  
**Clinical Effectiveness**  
Not examined in this study  

**Adverse Events**  
There was no statistically significant difference in rates of RE between ALIF with rhBMP-2 and ADR groups.  

$Retrograde ejaculation$ ($P = 0.7226$)  
ALIF with rhBMP-2 (n/N) = 7.4% (4/54)  
ADR (n/N) = 9.8% (4/41) |  
**Adverse Events**  
"This study found that RE occurred at a similar rate in patients treated with ADR and ALIF with BMP." (pp. 1785) |
| Rowan et al., 2012<sup>24</sup> | **Clinical Effectiveness**  
Not examined in this study |  
**Clinical Effectiveness**  
Not examined in this study  

**Adverse Events**  
There was no statistically significant differences in occurrence of postoperative leg pain in BMP-2 and control groups.  

$Postoperative leg pain$ (n/N) ($P = 0.140$)  
rhBMP-2 25% (16/64)  
control 12.5% (5/40) |  
**Adverse Events**  
"RhBMP-2 associated radiculitis presenting as immediate postoperative leg pain without MRI evidence of neuronal compression occurs in 17% of the patients with rhBMP-2 assisted fusion. Patients should be preoperatively counselled regarding immediate
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Main Study Findings</th>
<th>Author's Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tressler et al., 2011</strong>&lt;sup&gt;26&lt;/sup&gt;</td>
<td><strong>Clinical Effectiveness</strong>&lt;br&gt;Treatment of long bone nonunion&lt;br&gt;There were no statistically significant differences between rhBMP-2 and autograft groups in observed outcomes except for OR time and intraoperative blood loss.&lt;br&gt;<em>Operating time (minutes) (P = 0.0007)</em>&lt;br&gt;rhBMP-2 ± std dev = 168.9 ± 86.5&lt;br&gt;Autograft ± std dev = 257.9 ± 93.0&lt;br&gt;<em>Intraoperative blood loss (mL) (P = 0.01)</em>&lt;br&gt;rhBMP-2 ± std dev = 331.6 ± 357.2&lt;br&gt;Autograft ± std dev = 554.6 ± 447.8&lt;br&gt;<em>Proportion of healed nonunion (n/N) (P = 0.09)</em>&lt;br&gt;rhBMP-2 = 68.4% (13/19)&lt;br&gt;Autograft = 85.1% (63/74)</td>
<td>Clinical Effectiveness&lt;br&gt;&quot;These outcomes suggest that rhBMP-2 may provide a suitable alternative to autologous iliac bone graft, with the possible advantages of shorter operative time and reduced intraoperative blood loss, and may be considered as part of the orthopedic surgeon’s treatment options.&quot; (pp. 877)</td>
</tr>
<tr>
<td><strong>Gerszten et al., 2011</strong>&lt;sup&gt;19&lt;/sup&gt;</td>
<td><strong>Clinical Effectiveness</strong>&lt;br&gt;Overall there were no clinically and statistically significant differences in outcomes found between rhBMP-2 and autograft treatment groups.&lt;br&gt;<em>L5-S1 fusion rate with AxiaLIF (P = 0.269)</em>&lt;br&gt;rhBMP-2 = 96%&lt;br&gt;Autograft = 93%</td>
<td>Clinical Effectiveness&lt;br&gt;&quot;In our case-matched series, clinical outcomes were similar for patients who underwent an AxiaLIF L5–S1 interbody fusion with or without rhBMP-2. The data strongly suggest that there is a high confidence for no effect on fusion rate by using rhBMP-2.&quot; (pp. 1027)</td>
</tr>
</tbody>
</table>

**Adverse Events**

Tressler et al., 2011<br>No major adverse events reported.

Gerszten et al., 2011<br>No major complications identified in study.
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Main Study Findings</th>
<th>Author's Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carragee et al., 2011&lt;sup&gt;17&lt;/sup&gt;</td>
<td><strong>Clinical Effectiveness</strong>&lt;br&gt;Not examined in this study.</td>
<td><strong>Clinical Effectiveness</strong>&lt;br&gt;Not examined in this study.</td>
</tr>
<tr>
<td></td>
<td><strong>Adverse Events</strong>&lt;br&gt;Retrograde Ejaculation in ALIF ($P = 0.0025$)&lt;br&gt;rhBMP-2 (90%CI): 7.3%(2.11%, 12.39%)&lt;br&gt;Autograft (90%CI): 0.6%(0.37%, 1.51%)</td>
<td><strong>Adverse Events</strong>&lt;br&gt;“This study confirms previous reports of a higher rate of RE in ALIF procedures using rhBMP-2. This may be an important consideration in subjects concerned with sterility after surgery.” (pp. 511)</td>
</tr>
<tr>
<td>Williams et al., 2011&lt;sup&gt;27&lt;/sup&gt;</td>
<td><strong>Clinical Effectiveness</strong>&lt;br&gt;Not examined in this study.</td>
<td><strong>Clinical Effectiveness</strong>&lt;br&gt;Not examined in this study.</td>
</tr>
<tr>
<td></td>
<td><strong>Adverse Events</strong>&lt;br&gt;Excluding anterior cervical fusions no significant differences between fusions with and without BMP with regard to overall complications, wound infections or epidural hematomas/seromas.&lt;br&gt;Anterior cervical fusions&lt;br&gt;<strong>Overall complications (n/N) ($P &lt; 0.001$)</strong>&lt;br&gt;rhBMP-2 = 5.8% (38/652)&lt;br&gt;control = 2.4% (110/4532)&lt;br&gt;&lt;br&gt;<strong>Wound infection (n/N) ($P &lt; 0.001$)</strong>&lt;br&gt;rhBMP-2 = 2.1% (14/652)&lt;br&gt;control = 0.4% (17/4532)</td>
<td><strong>Adverse Events</strong>&lt;br&gt;“BMP use with anterior cervical fusion was associated with an increased incidence of complications. Use of BMP was not associated with more complications in thoracolumbar and posterior cervical fusions.” (pp. 1685)</td>
</tr>
<tr>
<td>Lee et al., 2010&lt;sup&gt;21&lt;/sup&gt;</td>
<td><strong>Clinical Effectiveness</strong>&lt;br&gt;PLF&lt;br&gt;Overall there were no other clinically and statistically significant difference in outcomes found between rhBMP-2 and autograft treatment groups aged 65 years and older.</td>
<td><strong>Clinical Effectiveness</strong>&lt;br&gt;“In patients 65 years and older, rhBMP-2 with allograft may lead to acceptable fusion rates and fusion times, good clinical outcomes and reduced perioperative complications.” (pp. 930)</td>
</tr>
<tr>
<td></td>
<td><strong>Adverse Events</strong>&lt;br&gt;Overall there were no clinically and statistically significant difference in adverse events found between rhBMP-2 and autograft treatment groups aged 65 years and older.</td>
<td>“... when compared to patients under 65 years of age undergoing posterolateral lumbar fusion, the use of rhBMP-2 [like autograft] was not sufficient to overcome all...”</td>
</tr>
<tr>
<td>First Author, Publication Year</td>
<td>Main Study Findings</td>
<td>Author's Conclusions</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Yaremchuk et al., 2010&lt;sup&gt;28&lt;/sup&gt;</td>
<td><strong>Clinical Effectiveness</strong>&lt;br&gt;Cervical spine surgeries&lt;br&gt;<em>Postoperative hospital stay (days)</em>&lt;br&gt;(P = 0.001)&lt;br&gt;rhBMP-2 ± std dev = 7.2 ± 11.1&lt;br&gt;Autograft ± std dev = 4.3 ± 5.2</td>
<td>aspects of the age-related weakened osteoinductive capacity encountered in elderly patients.” (pp. 930)</td>
</tr>
<tr>
<td></td>
<td><strong>Adverse Events</strong>&lt;br&gt;<em>Tracheotomies</em>&lt;br&gt;OR (95% CI) = 4.87 (1.23-19.23)&lt;br&gt;(P = 0.024)&lt;br&gt;rhBMP-2 (n/N) = (8/260)&lt;br&gt;control (n/N) = (3/515)</td>
<td>Clinical Effectiveness&lt;br&gt;Not examined in this study</td>
</tr>
<tr>
<td></td>
<td><em>Unplanned intubations after surgery</em> OR (95% CI) = 3.91 (1.60-9.54)&lt;br&gt;(P = 0.008)&lt;br&gt;rhBMP-2 (n/N) = (16/260)&lt;br&gt;control (n/N) = (8/515)</td>
<td>Adverse Events&lt;br&gt;“The increased incidence of unplanned intubations and tracheotomies demonstrates the risk associated with BMP in cervical spinal procedures.”&lt;br&gt;“With the results of this retrospective case study demonstrating increased morbidity and mortality associated with the use of BMP in the cervical spine, the off-label utilization of BMP in cervical spine procedures should be reconsidered.” (pp. 1954)</td>
</tr>
<tr>
<td></td>
<td><em>Readmissions</em> OR (95% CI) = 1.96 (1.03-3.70)&lt;br&gt;(P = 0.040)&lt;br&gt;rhBMP-2 (n/N) = (23/260)&lt;br&gt;control (n/N) = (26/515)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Dysphagia</em> OR (95% CI) = 8.94 (3.63-21.99)&lt;br&gt;(P = 0.001)&lt;br&gt;rhBMP-2 (n/N) = (18/260)&lt;br&gt;control (n/N) = (17/515)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Dysnea</em> OR (95% CI) = 2.43 (1.53-3.88)&lt;br&gt;(P = 0.001)&lt;br&gt;rhBMP-2 (n/N) = (53/260)&lt;br&gt;control (n/N) = (41/515)</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>Respiratory Failure</em> OR (95% CI) = 3.35 (1.88-5.97)&lt;br&gt;(P = 0.001)&lt;br&gt;rhBMP-2 (n/N) = (34/260)&lt;br&gt;control (n/N) = (24/515)</td>
<td></td>
</tr>
<tr>
<td>First Author, Publication Year</td>
<td>Main Study Findings</td>
<td>Author's Conclusions</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>---------------------</td>
<td>----------------------</td>
</tr>
</tbody>
</table>
| Taghavi et al., 2010<sup>25</sup> | **Clinical Effectiveness**  
PLF  
There were no statistically significant differences between rhBMP-2 and autograft groups except for the time to fusion overall and time to fusion for single-level fusion.  
*Time to fusion (single and multi-level) (days) (P = 0.033)*  
rhBMP-2 ± std dev = 218.4 ± 63.8  
Autograft ± std dev = 270.0 ± 60.4  
*Time to fusion (single-level) (days) (P = 0.001)*  
rhBMP-2 ± std dev = 199.8 ± 49.8  
Autograft ± std dev = 276.7 ± 29.8  
**Adverse Events**  
20% (4/20) patients in the autograft group complained of persistent donor-site pain at 2-year follow-up. | **Clinical Effectiveness**  
"rhBMP-2 may be an appropriate alternative to autogenous bone graft in both single- and multi-level revision PLF." (pp. 1144) |
| Rihn et al., 2009<sup>23</sup> | **Clinical Effectiveness**  
There was no other statistically significant difference in clinical effectiveness outcomes found between rhBMP-2 and autograft treatment groups.  
*Radiographic nonunion (P = 0.90)*  
TLIF with rhBMP-2 (n/N) = 3.5% (3/86)  
TLIF with autograft (n/N) = 3.0% (1/33)  
**Adverse Events**  
There was no statistically significant differences in adverse events including overall rate of complications except for complications that were unique to rhBMP-2 and autograft.  
*Unique complications for rhBMP-2 (n/N)*  
Lumbar seroma 1.2% (1/86)  
Ectopic bone formation 2.3% (2/86)  
Vertebral osteolysis 5.8% (5/86)  
Dural tear 4.7% (4/86)  
Malpositioned instrumentation 2.3% (2/86)  
Retained drain 1.2% (1/86)  
*Unique complications for Autograft (n/N)*  
Persistent donor-site pain 30.3% (10/33)  
Donor-site infection 3%(1/33) | **Clinical Effectiveness**  
"The procedure, whether performed using iliac crest autograft or rhBMP-2, provides over a 95% fusion." (pp. 628)  
**Adverse Events**  
"Although the use of rhBMP-2 eliminates autograft donor-site morbidity, additional complications such as postoperative radiculitis and ectopic bone are associated with its use. Further studies are needed to determine the mechanism of these complications and methods to minimize their occurrence." (pp. 628) |
<table>
<thead>
<tr>
<th>First Author, Publication Year</th>
<th>Main Study Findings</th>
<th>Author's Conclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACS=absorbable collagen sponge; ADR=artificial disc replacement; AIBG=autologous iliac crest bone graft; ALIF=anterior lumbar interbody fusion; AxiaLIF=axial lumbar interbody fusion; BMP=bone morphogenetic proteins; CI=confidence interval; DBM=demineralized bone matrix; DJD=degenerative disc disease; NRS=numerical rating scales; ODI=Oswestry Disability Index; OR=odds ratio; PCS=prospective cohort study; PLF=posterolateral lumbar fusion; PLIF=posterolateral lumbar interbody fusion; RCT=randomized controlled trial; RE=retrograde ejaculation; RR=relative risk; rhBMP=recombinant human bone morphogenetic protein; SF-36=36-Item Short Form Health Survey; TLIF=transforaminal lumbar interbody fusion</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>