NHS England

Evidence Review:
Bone morphogenetic protein-2 in spinal fusion

NHS England

Evidence Review:
Bone morphogenetic protein-2 in spinal fusion

First published:	December 2015
Updated:	Not applicable
Prepared by	Turnkey Clinical Evidence Review Team on behalf of NHS England Specialised Commissioning

Contents

1. Introduction

NHS England is responsible for commissioning complex spinal surgery as set out within the Manual for Prescribed Services (NHS England, November 2012). The commissioning criteria for complex spinal surgery which is commissioned by NHS England is documented within NHSE Policy D14/S/a.

Spinal fusion surgery permanently joins bones in the spine to ensure that there is no movement between them. The aim of a successful fusion is to reduce pain and disability. Fusions can be performed by removing the intervertebral disc and replacing it with a cage designed to maintain (or correct) the anatomical alignment of the lumbar spine. The cage is filled with material to encourage a fusion to occur.

Primary anterior lumbar surgery and revision surgery and posterior instrumented lumbar spinal surgery of more than 2 levels is commissioned by NHS England specialised commissioning teams.

A lumbar spinal fusion is performed (a) when the pain is thought to be due to degenerative change at one or two levels in the lumbar spine (b) to stabilise the spine following decompression of neurological structures where the decompression results in potential instability (c) to correct and stabilise a spinal deformity which is usually performed at multiple levels and may require decompression of the neurological structures.

The use of autologous bone graft (ABG), typically an iliac crest bone graft (ICBG), as an adjunct to spinal fusion surgery is considered the gold standard. Whilst the use of bone graft possesses the three key properties required for bone formation: osteoconductivity (acts as a scaffold allowing native bone to perpetuate), osteoinductivity (stimulates osteoprogenitor cells to differentiate into osteoblasts that then begin new bone formation) and osteogenicity (osteoblasts originating from the bone graft material contribute to new bone growth along with bone growth generated via the other two mechanisms) it may not be suitable for all patients, especially those who do not have sufficient quality iliac of crest bone material, where it has been harvested for previous surgery or where the bone is required for secure fixation as part of the spinal instrumentation.

Bone morphogenetic protein (BMP) is a graft substitute. Currently, the BMP with the widest clinical application is recombinant human bone morphogenetic protein-2 (rhBMP-2), an osteoinductive bone growth factor that is a member of the transforming growth factor-b superfamily.

2. Summary of results

The evidence review has sought to establish the clinical effectiveness, safety and cost effectiveness of rhBMP-2 in comparison with iliac crest bone graft for anterior lumber spinal fusion surgery and posterior instrumented spinal surgery to inform the NHS England policy.

Clinical effectiveness:

The evidence for clinical effectiveness of BMP is based on five good quality independent systematic reviews and meta-analyses (Chen et al., 2012; Fu et al., 2013; Simmonds et al., 2013; Zhang et al., 2014; Noshchenko et al., 2014). The number of studies included in the reviews varied depending on the inclusion and exclusion criteria but all included 8 RCTs evaluating rhBMP-2 with ICBG for lumbar fusion (including anterior lumber spinal fusion). All reviews compared rhBMP-2 with ICBG for spinal fusion and the primary outcomes were rate of fusion and improvement of clinical symptoms based on the ODI and the SF-36, physical scale. The quality of reporting secondary outcomes varied across studies.

Fu et al. (2013) and Simmonds et al. (2013) systematic reviews were conducted as part of The Yale University Open Data Access (YODA) Project. In addition to the published studies, individual-participant data was obtained from sponsors or investigators to address the issue of publication bias.

The results of the analysis on the primary outcome measure indicate that compared with ICBG, rhBMP-2 in lumbar fusion (single level anterior or posterior fusion) has higher rates of radiographic fusion at 2 years follow up period. The Relative Risk (RR) for radiographic fusion varied from 1.13 to 1.19, with 2 reviews showing a statistically significant difference.

Successful fusion was not, however, correlated with improvement in clinical outcomes as measured by: the Oswestry Disability Index (ODI), return to work, back pain, leg pain and SF-36. Both groups had significant

improvements in clinical outcomes but at 2 years follow up there was no statistically significant difference between the two groups. Similar results were observed in a recently published RCT of 197 patients with a 4 years follow up (Hurlbert et al., 2013). After 4 years of follow up, radiographical fusion rates remained significantly higher in patients treated with rhBMP-2 (94%) than those who received autograft (69%) (P = 0.007). However, SF-36, ODI and leg/back pain scores were comparable between the 2 groups.

The rate of non-union at 2 years postoperative was significantly lower in the rhBMP-2 groups (including off-label use) and was approximately half that of the ICBG groups. However, this did not lead to similar improvement for patient centred outcomes and funnel plot analysis indicated an asymmetry of published results, with a tendency to underestimate the non-union risk for rhBMP-2, this may be suggestive of a publication bias (Noshchenko et al., 2014).

Subgroup analysis by type of surgery: anterior lumbar spine (ALIF) and posterior lumbar spine (PLF or PLIF) found similar results for fusion rates and clinical outcomes (Fu et al., 2013).

Radiological fusion and patient related clinical outcomes:

As radiological fusion is used as the primary outcome measure, the clinical relevance of successful fusion after lumbar arthrodesis with rhBMP-2 or ICBG was studied in a meta-analysis by Noshchenko et al. (2015). This study concluded that patients who had radiological fusion had significantly better clinical outcome measures (ODI and Numeric Rating Scales (NRS) for back and leg pain) but fusion used on its own was a poor predictor of clinical outcomes, indicating that other factors contributed to patient related clinical outcome measures.

Overall, it can be concluded that successful fusion using rhBMP-2 is not strongly correlated with improvement in clinical outcomes and it should be noted that no trials were independent of industry sponsorship.

Safety:

The initial reports from industry sponsored trials reported low levels of side effects resulting from the use of rhBMP-2. However, a systematic review by Carragee et al. (2011) reported that adverse events associated with rhBMP-2 use in spine fusion ranged from 10% to 50% (depending on approach) in comparison to the 0% reported in some industry sponsored trials.

Adverse events for ALIF were not directly reported however anterior cervical fusion with rhBMP-2 has an estimated 40% greater risk of adverse events in the early postoperative period, including life-threatening events. Posterior lumbar interbody fusion (PLIF) use was associated with radiculitis, ectopic bone formation, osteolysis, and poorer global outcomes. In posterolateral fusions, the risk of adverse effects associated with rhBMP-2 use was equivalent to, or greater than, that of iliac crest bone graft harvesting, and 15% to 20% of subjects reported early adverse events of back pain and leg pain. Higher doses of rhBMP-2 were also associated with a greater apparent risk of new malignancy (Carragee et al., 2011).

Similar levels of side effects from rhBMP-2 have been reported in other reviews. A meta-analysis, involving 184,324 patients (28,815 rhBMP-2 group, 155,509 ICBG group) from 26 studies published between 2002-2013 by Vavken et al. (2015), reported significantly higher risk of general complications with rhBMP-2 compared to iliac crest bone graft (ICBG) with an odds ratio (OR) of 1.78 (95% CI 1.20–2.63, p = 0.004). The OR for heterotrophic ossification (HO) was 5.57 (95% CI 1.90–16.36, p = 0.002), for retrograde ejaculation 3.31 (95% CI 1.20–9.09, p = 0.020), and for cervical swelling 4.72 (95%CI 1.42–15.67, p = 0.011), all significantly higher in the rhBMP-2 group. Other outcomes such as perioperative clinical outcomes including blood loss, complications/adverse events, and hospital stay were not significantly different between the rhBMP-2 and ICBG groups.

A recent study retrospectively analysed data from 460,773 patients who underwent lumbar spine fusion either without rhBMP-2 (69.3%) or with (30.7%) (Savage et al, 2015). A slightly lower complication rate was reported with rhBMP-2 group (18.2%) compared to the control group (18.7%). This difference did not appear to be very significant (Relative Risk (RR) 0.976 (CI 0.963–0.989) ($p < 0.001$). In both treatment groups, patients older than 65 years had a significantly higher risk of postoperative complications than the younger patients ($p < 0.001$). However in patients younger than 65 years, those treated with rhBMP-2 had higher rate of complications compared to control group (Relative Risk (RR)1.042 (CI 1.017–1.067, $p<0.001$), whereas in the patients ≥ 65 years old, the opposite was true i.e. lower complication rates in rhBMP-2 group (Relative Risk (RR) 0.950 (CI 0.935–0.065). For both males and females, the complication rates were lower in the rhBMP-2 group than in the

control group but it was only significantly lower in females (Relative Risk (RR) of 0.974 [CI 0.953–0.995,

p=0.015] in males and 0.976 [CI 0.960–0.993, p=0.005] in females). The authors also report 90-day reoperation rates of 1.84% in the control group, which was significantly lower compared to 2.03% in the rhBMP-2 group (Relative Risk (RR) 1.108 (CI 1.060–1.158, p<0.01). In both the control and rhBMP-2 groups, patients younger than 65 years were more likely to have a reoperation than patients older than 65 years (p < 0.001). Although this is a large study the difference in response (overall, age, and gender specific) for rhBMP-2 and non-rhBMP-2 patients cited by the authors has limited implication in a real world setting given the nearly 1 relative risk in all cases.

The outcomes that favoured rhBMP-2 compared to ICBG were mean operative time for patients, which was significantly less for patients treated with rhBMP-2 than that of patients who underwent ICBG harvest, and the number of patients requiring additional surgical treatment during 2 postoperative years, which was also significantly lower in the rhBMP-2 groups (Zhang et al., 2014).

Nearly 50% of the patients who underwent lumbar fusion with ICBG experienced donor site pain at 2 years follow up and the risk of complications at the ICBG donor site was 7% (Noshchenko et al., 2014).

Cost effectiveness:

The evidence of cost effectiveness is based on two studies, one systematic review of studies evaluating cost effectiveness of rhBMP-2 against ICBG (Hsu et al., 2014) and one cost utility analysis in 33 patients receiving posterior lumbar fusion using rhBMP-2 (Alvin et al., 2014).

The systematic review included 5 studies (Polly et al., 2003; Garrison et al., 2007; Alt et al., 2009; Carreon et al., 2009; AHRQ, 2010) that compared fusion with rhBMP-2 to fusion with ICBG in patients with degenerative disease of the lumbar spine. In all cases, 2 year time horizon was used and no discounting was performed. All relied on a single non inferiority randomized trial (Burkus et al., 2002) for clinical data that served as the pivotal trial for FDA approval of Medtronic Sofamor Danek Inc., (Memphis, TN) Infuse (rhBMP-2). Two studies (AHRQ, 2010; Garrison et al., 2007) relied solely on this RCT, one (Alt et al., 2009) also used data from 2 other nonrandomized trials of the same grafts inserted laparoscopically, and one (Polly et al., 2003) also used expert opinion. Two studies (AHRQ, 2010; Garrison et al., 2007) undertook cost-utility analyses (CUA) from a payer perspective. Both derived utility estimates from unpublished preoperative and 6-month SF-36 data from the trial.

There were conflicting conclusions reached depending on the type of data used, cost-measurement methods and study design. For example, the National Health Service study used cost of treatment and hospitalization data from the United Kingdom and concluded that rhBMP-2 was not cost-effective. rhBMP-2 versus ICBG was associated with £120,390 per QALY gained. No sensitivity analysis was performed.

Conversely, Alt et al. (2009) reported data including return-to-work parameters from 3 different European countries and concluded that the increased loss of productivity seen from the ICBG group resulted in a savings with use of rhBMP-2 per patient. Outcome measures used in the analysis included need for secondary surgery and return-to-work. Compared with ICBG, rhBMP-2 use resulted in savings ranging from £236 to £529 per patient as a result of decreased rates of secondary surgery and £4938 to £5450 savings from prevented lost productivity. The authors concluded that from a societal perspective, use of rhBMP-2 resulted in savings over time that offset the higher upfront cost of rhBMP-2 use compared with ICBG. All of the studies in the review had limitations including: lack of time horizon discounting, basis on a single RCT with a short time scale (2 years), lack of sensitivity analysis (Alt et al., 2009; Carreon et al., 2009) and no inclusion of indirect costs in all except Alt et al. (2009). All studies except AHRQ (2010) and Garrison et al. (2007) were linked to sponsoring from manufacturers of rhBMP-2. In another study, Alvin et al. (2014) demonstrated that the 1-year cost-utility ratio (Total Cost/ΔQALY) for the ICBG cohort was significantly lower (£94,177/QALY gained) than that of the rhBMP-2 cohort (£179,092/QALY gained) (P<0.01).

A cost effective analysis by Virk et al. (2012) suggested that while rhBMP-2 has better cost per QALY (£10,910/QALY) compared to ICBG (£14,008/QALY), the sensitivity analysis shows that rhBMP2 is not the most cost-effective option if the revision rate is significantly raised. This is significant considering that the findings from a recent population level study by Savage et al. (2015) showed that the 90 day reoperation rate in a group using rhBMP-2 for lumbar spinal fusion was significantly higher than group using non- rhBMP-2 methods (RR1.108, CI 1.060–1.158).

Based on the current evidence it can be concluded that there is no clear evidence that using rhBMP-2 is more cost effective than ICBG. If anything, the evidence suggests that the cost per QALY of rhBMP-2 is higher than ICBG but this is based on studies with low levels of evidence and study design, and industry sponsorship.

[Original figures provided in euros and US dollars were converted to the nearest full pound based on conversion rate on 17/11/2015 of £1 to 1.43 euro and £1 to $1.52 and is provided as a guideline for comparison only]

This clinical evidence review also considered the following specific questions related to the clinical effectiveness, safety and cost effectiveness of bone morphogenetic protein-2 (rhBMP-2)

Question 1: Is the use of rhBMP-2 safe and effective (in terms of clinical and radiographical outcomes) when used in adults for revision spinal fusion surgery when autologous bone graft (ABG) has previously been used and failed to achieve union (pseudoarthrosis)?

Question 2: Is the use of rhBMP-2 safe and effective (in terms of clinical and radiographical outcomes) when used in adults for primary spinal fusion surgery where there is high risk of pseudoarthrosis compared with autologous bone graft (ABG) alone?

Response to question 1:

Evidence on the use of rhBMP-2 in revision spinal fusion surgery is available from one retrospective cohort study by Taghavi et al. (2010), however this considered posterior lumbar fusion only. The objective of this study was to determine the efficacy of rhBMP-2 or local bone, to either allograft combined with bone marrow aspirate (BMA) or autograft, in revision instrumented, posterolateral fusions (PLF). Indications for revision surgery included: symptomatic pseudoarthrosis (pain and/or instability) following a previous PLF for degenerative conditions of the lumbar spine, such as degenerative disc disease, stenosis, or spondylolisthesis. Sixty-two patients were divided into 3 groups: Group 1 (n = 24) received rhBMP-2, Group 2 (n =18) received BMA/allograft, and Group 3 (n = 20) received autograft. The exact source of autograft bone for Group 3 was not clearly defined. All 3 cohorts received supplemental local bone. Static and dynamic radiographs were used to assess fusion and clinical outcome was determined through Visual Analogue Scale (VAS) scores. At 2 years follow up, there was no difference between group 1 and 3, a fusion rate of 100% was observed for both groups. Similarly, no difference in VAS score was observed between group 1 and group 3.

The ability to generalise the results is limited due to the retrospective nature of the study design and small sample size.

Dorward et al. (2013) evaluated cervical fusion rates with rhBMP-2 in 57 patients, this group included 48 patients (84.2%) who had undergone previous cervical surgery, and 42.1% who had a pre-existing non-union. Successful fusion was seen in 89.5% of patients. The neurologic symptoms were resolved postoperatively in 50 patients (87.7%) and both VAS and Neck Disability Index (NDI) scores improved significantly from baseline. The results were not provided in subgroups by previous surgery or non-union. The study is also limited by the lack of a comparator group, a lack of randomisation and small sample size.

Response to question 2:

There are a limited number of studies evaluating the risk of pseudoarthrosis when using rhBMP-2 in people with one or more risk factors.

A study by Lee et al. (2013) compared fusion rates for rhBMP-2 versus autograft in patients with fusion-related risk factors. Fusion related high risk factors were defined as i) old age (>65 years) ii) pseudoarthrosis with a T-score of less than -2.5 based on dual energy X-ray absorptiometry iii) those who had continuously smoked for at least 1 year before surgery (iv) postoperative, medical comorbidities, including those who were receiving treatment for 2 or more concurrent medical diseases such as diabetes mellitus, hypertension, and thyroid disease v) revision surgery including cases in which surgery was performed for pseudoarthrosis, or vi) multilevel fusion cases in which >2 levels were surgically treated. One hundred and ninety-five patients were divided into 4 groups depending on fusion material and the presence/absence of fusion-related risk factors for non-union; Group A was defined as rhBMP-2 used in the presence of high-risk factors (FRRF), group B was defined as rhBMP-2 used in the absence of FRRF, group C was defined as autograft used in the presence of FRRF and group D was defined as autograft used in the absence of FRRF.

Although time to fusion was faster in group A than in group C in all fusion-related risk factors (age, sex, revision, fusion level, smoking, DM, osteoporosis, and comorbidity), there was no statistically significant difference between groups A and C at 2 years follow up. Similarly, fusion rate was higher in group A than in group C in other fusion related risk factors, except revision surgery but there was no statistically significant difference between groups A and C in all fusion-related risk factors.

There was no significant difference in results for subjects who were over 65 years of age or for smokers.

3. Research questions

1. Is the use of rhBMP-2 safe and effective (in terms of clinical and radiographical outcomes) when used in adults for revision spinal fusion surgery when autologous bone graft (ABG) has previously been used and failed to achieve union (pseudoarthrosis)?
2. Is the use of rhBMP-2 safe and effective (in terms of clinical and radiographical outcomes) when used in adults for primary spinal fusion surgery where there is high risk of pseudoarthrosis compared with autologous bone graft (ABG) alone?

4. Methodology

A review of published, peer reviewed literature has been undertaken based on the research questions set out in Section 3 and a search strategy agreed with the lead clinician and public health lead for this policy area. This has involved a PubMed search and search of the Cochrane database for systematic reviews, in addition to review of any existing NICE or SIGN guidance. The evidence review has been independently quality assured.

An audit trail has been maintained of papers excluded from the review on the basis of the inclusion and exclusion criteria agreed within the search strategy. The full list has been made available to the clinicians developing the policy where requested.

5. Results

A detailed breakdown of the evidence is included in the Appendix.

Appendix 1

Level	Study design and intervention				Outcomes						Reference	Other		
Level of evidence	Study design	Study size	Intervention	Category	Primary Outcome	Primary Result	Secondary Outcome	Secondary Result	Study Endpoint	Study Endpoint Result	Reference	Complications noted	Benefits noted	Comments
1+	Systematic	Of the 5 studies, 3 studies are based on one RCT with 279 subject, one is based on 1 RCT and 2 prospective studies and 6th is based on a RCT with 106 patients	BMP	Cost effectiveness	Cost per QALY, Incremental cost per QALY per year, reduces costs from societal perspective calculated as reduced need for secondary surgery, and average prevented lost productivity	Among the 5 studies that met the inclusion criteria to compare cost-effectiveness of rhBMP-2 versus ICBG, discordant conclusions were reached depending on the type of data used, cost-measurement methods, and study design. For example, the National Health Service study used cost of treatment and hospitalisation data from the United Kingdom and concluded that rhBMP-2 was not cost-effective. Conversely, Alt et al reported data including return-to-work parameters from 3 different European countries and concluded that the increased loss of productivity seen from the ICBG group resulted in a savings with use of BMP per patient. Finally, the AHRQstudy used Centers for Medicare & Medicaid Services cost data for initial and secondary interventions and concluded that rhBMP-2 would be cost-effective compared with ICBG if there was no additional cost for the product; however, at a price of $3000, there was no significant difference.	None	-	-	-	Hsu, Wellington K.; Hashimoto, Robin E.; Berven, Sigurd H.; Nassr, Ahmad. Biological substitutes/extenders for spinal arthrodesis: which agents are cost-effective?. Spine. 2014,	-	As in primary outcome measures	Population: Lumbar spine 1-level. DDD and disabling symptoms for ≥ 6months. 42.5+ years. Comments: This systematic review of cost effectiveness was designed to evaluate: 1. Is there evidence to suggest that the use of rhBMP-2 for spinal arthrodesis is cost-effective compared with the use of iliac crest bone graft (ICBG)? 2. Is there evidence to suggest that the use of allograft or bone graft extenders (demineralised bone matrices, synthetic carriers, and allogeneic stem cell products) is cost-effective in spinal arthrodesis compared with the use of ICBG? 3. Are there differences in cost-effectiveness with the use of allograft bone versus polyetheretherketone cages with bone graft substitutes in anterior cervical discectomy and fusion (ACDF). There were 5 studies that met the inclusion criteria to compare cost-effectiveness of Q1 rhBMP-2 versus ICBG (AHRQ 2010; Alt et al., 2009; Carreon et al., 2009; Garrison et al., 2007; Polly et al., 2003). There were discordant conclusions reached depending on the type of data used, cost-measurement methods, and study design. For example, the National Health Service study used cost of treatment and hospitalisation data from the UK and concluded that rhBMP-2 was not cost-effective. Anterior lumbar interbody fusion with rhBMP-2 versus ICBG was associated with £120,390 per QALY gained, and thus is not likely to be cost effective, depending on the willingness to pay threshold. No sensitivity analysis was performed. Conversely, Alt et al. (2009) reported data including return-to-work parameters from 3 different European countries and concluded that the increased loss of productivity seen from the ICBG group resulted in a savings with use of BMP per patient. Outcome measures used in the analysis included need for secondary surgery and return-to-work. Costs included in the analysis were those of the initial treatment, secondary surgery, and cost of lost productivity and were reported in 2008 Euros. Compared with ICBG, rhBMP-2 use resulted in savings in each country as a result of decreased rates of secondary surgery; savings were in the range from €337 to €756 per patient. Because patients who received rhBMP-2 returned to work on average 43 days sooner than those who received ICBG, rhBMP-2 use was associated with €7052 to €7783 savings from prevented lost productivity. Overall, the savings from avoiding lost productivity accounted for 82% to 92% of the overall savings associated with rhBMP-2 use. The authors concluded that, from a societal perspective, use of rhBMP-2 resulted in savings over time that offset the higher upfront cost of rhBMP-2 use compared with ICBG. However, all studies included in the review had limitations including: no time horizon discounting, based on single RCT, short time scale (2 yrs), lack of sensitivity analysis (Alt et al,2009; Carreon et al, 2009), lack of inclusion of indirect costs in all except Alt et al (2009), models described in the paper not available for detailed review in two studies (Garrison et al., 2007 and Polly et al.2003). All studies except AHRQ (2010) and Garrison et al. (2007) were linked to sponsoring from manufacturers of BMP.

1++	Systematic	1138	Bone Morphogenetic Protein	Clinical effectiveness of the intervention	Patient-centred clinical evaluations before and 24 months after surgery, including: Oswestry Disability Index (ODI), SF-36, back and leg pain numeric or analog scales, rate of patients satisfied with surgical treatment outcome. (2) Perioperative outcomes: operative time (min), blood loss (mL), hospital stay (day), complications/ adverse events (rate).	The pooled 2-year postoperative clinical outcomes were equivalent in BMP and ICBG groups, and exceeded minimum clinically important differences for Oswestry Disability Index, SF-36 (physical scale), and numeric rating scale (back pain). ICBG was associated with increased pain and complications at the donor site (P<0.01). The pooled average operative time was 21 minutes less in BMP versus ICBG (P<0.001). The pooled rate of additional surgical treatment was 2 times less in the BMP than in the ICBG groups (P=0.006). The pooled risk of non-union at 24-month follow-up was 2 times less in the BMP than in the ICBG groups (P=0.037), however, this effect was likely biased.	Complications requiring additional surgical treatment during the 24 months follow-up. (2) Non-union (fusion failure) at 24-month follow-up. (3) Donor site complications. (4) Donor site pain at 24-month follow-up. (5) Wound infection. (6) Employment before and after surgery. (7) Bone/graft resorption. (8) Ectopic or heterotopic bone formation. (9) Retrograde ejaculation. (10) Postoperative back/leg pain or radiculopathy. (11) Postoperative cancer risk.	-	-	-	Noshchenko, Andriy; Hoffecker, Lilian; Lindley, Emily M.; Burger, Evalina L.; Cain, Christopher M. J.; Patel, Vikas V.. Perioperative and long-term clinical outcomes for bone morphogenetic protein versus iliac crest bone graft for lumbar fusion in degenerative disk disease: systematic review with meta-analysis. J Spinal Disord Tech. 2014	-	-	Population: Degenerative disk disease. Varying ages. Comments: This is an independently conducted, well designed and presented meta-analysis of studies comparing rhBMP-2 with ICBG . The study has clear eligibility criteria with inclusion and exclusion criteria, primary and secondary outcomes. The risk of bias assessment and evaluation of validity was done using the 12 criteria recommended by the Cochrane review group. Data extraction and assessment of clinical relevance was done by two reviewers independently. Robust statistical methods were used to assess heterogeneity, clinical relevance was assessed using questions recommended by Cochrane Back Review Group. A random effect model was used in the meta-analyis and funnel plots and statistics tests (Egger's and Begg's) were done to assess publication bias. The pooled 2-year postoperative clinical outcomes were equivalent in the rhBMP-2 and ICBG groups, and exceeded minimum clinically important differences for ODI, SF-36 (physical scale), and numeric rating scale (back pain). ICBG was associated with increased pain and complications at the donor site (P<0.01). The pooled average operative time was 21 minutes less in rhBMP-2 versus ICBG (P<0.001). The pooled rate of additional surgical treatment was 2 times less in the rhBMP-2 than in the ICBG groups (P=0.006). The pooled risk of non-union at 24-month follow-up was 2 times less in the rhBMP-2 than in the ICBG groups (P=0.037). Some of the limitations of this review, as recognised by the authors, include: low quality of studies included (as evaluated using Cochrane review group quality assessment tool), lack of double blinding and the majority of the trials were sponsored by the manufacturers.
0	RCT	140	C1-C2 fusion with iliac crest bone graft (ICBG)	Clinical effectiveness of the intervention	Operative time, estimated blood loss, hospital length of stay, clinical results, perioperative complications, fusion rate, fusion time, and revision rate.	The fusion rate was 82.4% (56 of 68) in the rhBMP-2/ACS group and 78.7% (52 of 66) in the ICBG group (P=.782). Mean fusion time was 11 days shorter in the rhBMP-2/ACS group (81.8±29.4 days) than in the ICBG group (92.9±23.7 days) (P=.034). There were more wound complications requiring treatment in the rhBMP-2/ACS group (n=6; 8.8%) than in the ICBG group (n=2; 3.0%), although this was not statistically significant (P=.118).	None	-	-	-	Yan, Liang; Chang, Zhen; He, Baorong; Liu, Tuanjiang; Wang, Xiaodong; Guo, Hua; Hao, Dingjun. Efficacy of rhBMP-2 versus iliac crest bone graft for posterior C1-C2 fusion in patients older than 60 years. Orthopedics. 2014	There were more wound complications requiring treatment in the rhBMP-2/ACS group (n=6; 8.8%) than in the ICBG group (n=2; 3.0%), although this was not statistically significant (P=.118).	As in primary outcome measure	Population: C1-C2 fusion for age related cervical spine conditions. 60+ years. Comments: This study is published as and RCT however no full text was available from the publishers so it is difficult to comment on the generalisability of the results due to lack of info on study methodology and analysis. However, the results suggest that there was no significant difference in the fusion rate of C1 C2 vertebrae but the group using rhBMP-2+ICBG had shorter healing time and higher rate of wound complication requiring treatment.

1++	Systematic	1852	rhBMP	Clinical effectiveness of the intervention	(1) solid fusion rate, (2) clinical outcomes, (3) complications, and (4) the reoperation rate.	BMPs group significantly increased fusion rate (RR: 1.13; 95% CI 1.05–1.23, P = 0.001), while there was no statistical difference in overall success of clinical outcomes (RR: 1.04; 95% CI 0.95–1.13, P = 0.38) and complications (RR: 0.96; 95% CI 0.85–1.09, p = 0.54). A significant reduction of the reoperation rate was found in BMPs group (RR: 0.57; 95% CI 0.42–0.77, p = 0.0002). Subgroup analysis of the fusion rate stratified by the two types of BMPs yielded different results. Compared with ICBG, the use of BMP-2 can increase solid fusion rate, by contrast, pooled BMP-7 studies do not have similar effects.	(1) the operation time and blood loss, and hospital stay, (2) patient satisfaction with the treatment, (3) work status and return to work rate.	BMPs group had a significantly lower operating time (MD-0.32; 95% CI20.55, 20.08; P = 0.009), but no significant difference was found in the blood loss, the hospital stay, patient satisfaction, and work status.	-	-	Zhang, Haifei; Wang, Feng; Ding, Lin; Zhang, Zhiyu; Sun, Deri; Feng, Xinmin; An, Jiuli; Zhu, Yue. A meta analysis of lumbar spinal fusion surgery using bone morphogenetic proteins and autologous iliac crest bone graft. PLoS ONE. 2014	As in secondary outcome results	-	Population: Adult patients suffering from degenerative conditions of the lumbar spine requiring fusion; the main intervention was lumbar fusion. rhBMP-2 and rhBMP-7. Mean age of studies ranged from 40 - 73 yrs with majority of studies in the 40 - 50 years age range. Comments: This is an independently conducted, well designed and presented metaanalysis of studies comparing BMP with ICBG published between 2000-Nov 2013. The study has clear eligibility criteria with inclusion and exclusion criteria, primary and secondary outcomes. The risk of bias assessment and evaluation of validity was done using the 12 criteria recommended by the Cochrane review group. Data extraction and assessment of clinical relevance was done two reviewers independently. Robust statistical methods were used to assess heterogeneity, clinical relevance was assessed using questions recommended by Cochrane Back Review Group. A random effect model was used in the metaanalysis and Funnel plots and statistics tests (Egger's and Beggs's test were done to assess publication bias. Quasi-RCT and non-RCTs were excluded from the study and the quality of the overall evidence for each individual outcome was evaluated using GRADE system. The results of the metaanalysis showed that compared with ICBG, BMPs in lumbar fusion can increase the fusion rate, while reduce the reoperation rate and operating time. However, it doesn't increase the complication rate, the amount of blood loss and hospital stay. No significant difference was found in the overall success of clinical outcome of the two groups. However some limitation of this review as recognised by the authors including low quality of studies included as evaluated using Cochrane review group quality assessment tool, lack of double blinding and majority of the trails were sponsored by the manufacturers. Also as the search was restricted to reports of published RCTs only this could have resulted inclusion of studies with positive and statistically significant results only.
1+	RCT	197	Recombinant human bone morphogenetic protein (rhBMP-2)	Clinical effectiveness of the intervention	Radiographical fusion rate and clinical outcome for surgical lumbar arthrodesis	After 4 years of follow-up, radiographical fusion rates remained significantly higher in patients treated with rhBMP-2 (94%) than those who received autograft (69%) (P = 0.007). However 36-Item Short Form Health Survey, Oswestry Disability Index, and leg/back pain scores were comparable between the 2 groups. ICBG group experienced signifi cantly more graft-site complications.	-	-	-	-	Hurlbert, R. John; Alexander, David; Bailey, Stewart; Mahood, James; Abraham, Ed; McBroom, Robert; Jodoin, Alain; Fisher, Charles. rhBMP-2 for posterolateral instrumented lumbar fusion: a multicenter prospective randomized controlled trial. Spine. 2013	-	-	Population: Degenerative disc disease with corresponding back and/ or leg pain necessitating instrumented fusion. Failure of conservative (nonsurgical) care for at least 6 months. Adults, 18+ years. Comments: A RCT comparing ICBG with rhBMP-2 with main primary outcome being radiological fusion rate. After 4 years of follow-up, radiographical fusion rates remained significantly higher in patients treated with rhBMP-2 (94%) than those who received autograft (69%) (P = 0.007). However SF-36, ODI, and leg/back pain scores were comparable between the 2 groups. The ICBG group experienced significantly more graft-site complications. Main limitations of the study are: lack of clear methodology on randomisation, only mentions patients were randomised to either of the groups at the time of surgery, open label study, only the radiographers who interpreted the x-rays were blinded. Lack of blinding likely to create bias for patient reported outcomes including SF36, ODI. This study was also funded by the manufacturer- Medtronic.

1++	Systematic	Anterior Lunmbar Interbody fusion (ALIF) 465 patients from 5 RCTS. Posterior Lumbar Fusion 722 patients from 5 RCTs, PLIF-67 from 1 RCT	rhBMP2	Clinical effectiveness of the intervention	Fusion rates, Adverse events including retrograde ejaculation, urine retention, subsidence, ectopic bone formation, cancer and death	ALIF: Moderate-strength evidence of no consistent differences between rhBMP-2 and ICBG in overall success, fusion rates, or other effectiveness measures from 6 weeks through 24 months after surgery. The rhBMP-2 group had 3 points higher than ICBG group for SF-36 score at 3, 6, 12, and 24 months but not statistically different. At 24 months, fusion rates ranged from 60% - 100%, and the average overall success rate was 61% for the rhBMP-2 group and 53% for the ICBG group and was not statistically different. Adverse events were common. Metaanalysis showed no significant differences between groups for any specific adverse event, including lumbar radiculitis, although estimates were frequently imprecise, precluding strong conclusions. For retrograde ejaculation, subsidence (defined as sinking or settling of the device into bone), and urogenital problems, risk estimates favoured ICBG but the differences were not statistically significant. PLF: There was moderate-strength evidence of no consistent difference between rhBMP-2 and ICBG in effectiveness outcomes through 24 months. The fusion rate at 24 months ranged from 70% to 90% in the ICBG group and 86% to 100% in the rhBMP-2 group; the rate of overall success ranged from 40% to 60% in both groups. No significant difference were observed between the rhBMP-2 and ICBG groups in adverse events. But authors comment that estimates were frequently imprecise, precluding strong conclusions. The only exception was that the rhBMP-2 group had increased risk for back and leg pain through 4 weeks, although events unrelated to fusion surgery could not be not ruled out. Cervical Spine Fusion: Based on small study (n=33) rhBMP-2 and ICBG did not differ in effectiveness end points. rhBMP-2 was associated with a greater risk for adverse events than ICBG at 24 months (45 adverse events in 18 patients vs. 13 adverse events in 15 patients; rate ratio, 2.88 [CI, 1.30 to 6.41]). Evidence from observational studies suggest that rhBMP-2 was associated with significant increased risk for complications (odds ratio, 1.43 [CI, 1.12 to 1.70]), dysphagia or dysphonia (odds ratio, 1.63 [CI, 1.30 to 2.05]), and wound complications (odds ratio, 1.67 [CI, 1.10 to 2.53]). In posterior cervical spine fusion, there were no controlled trials of rhBMP-2 and 1 cohort study showed no difference in rates of major complications.	None	-	-	-	Fu, Rongwei; Selph, Shelley; McDonagh, Marian; Peterson, Kimberly; Tiwari, Arpita; Chou, Roger; Helfand, Mark. Effectiveness and harms of recombinant human bone morphogenetic protein-2 in spine fusion: a systematic review and meta-analysis. Ann. Intern. Med.. 2013	-	-	Population: Adult patients with chronic low back pain and a lumbar disc degenerative condition. Mean age varied form 40- 57.6 years. Anterior Lumbar Interbody Fusion (ALIF), Posterior Lumbar body Fusion (PLF), Posterior Lumbar Interbody Fusion (PLIF) and cervical spine fusion. Comments: A large scale, independent meta-analysis based on individual participant data obtained from the manufacturer and published data conducted as part of The Yale University Open Data Access (YODA) Project. The paper included sound methodology including: search strategy, inclusion criteria, stattical analysis for heterogeneity, bias and quality of studies and presenetaion of data as CI. Although the rhBMP-2 had higher rates of fusion at 24 months for both ALIF and PLF, it was not statistically different from ICBG. Similarly, there was no difference bewteen the two group for patient related factors measured using ODI and SF36 . The adverse events were common. Meta-analysis showed no significant differences between groups for any specific adverse event, including lumbar radiculitis, although estimates were frequently imprecise, precluding strong conclusions. For retrograde ejaculation, subsidence (defined as sinking or settling of the device into bone), and urogenital problems, risk estimates favoured ICBG but the differences were not statistically significant. The results of the study are generalisable and few limitations of the meta-analysis, as noted by the authors, include lack of blinding for outcome assessment and ascertaintment of harms was poor in trials. All studies included had links to industry.

1++	Systematic	1408	rhBMP2	Clinical effectiveness of the intervention	In addition fusion primary outcomes were those likely to be important to patients including The Oswestry Disability Index (ODI) and Neck Disability Index for cervical spinal surgery measure lower back and neck pain, respectively, on a scale from 0% (no pain) to 100% (extreme pain). TheShort Form-36 (SF-36) Physical Component Summary(PCS) assesses pain and physical function on a scale from 0% (worst) to 100% (best).	At 24 months, pain assessed by ODI scores were 3.5% lower (better) with rhBMP-2 than with ICBG (95% CI, 0.5% to 6.5%) and radiographic fusion was 12% higher (CI, 2% to 23%). At or shortly after surgery, pain was more common with rhBMP-2 (odds ratio, 1.78 [CI, 1.06 to 2.95]). Cancer was more common after rhBMP-2 (relative risk, 1.98 [CI, 0.86 to 4.54]), but the small number of events precluded definite conclusions.A subgroup analyses to investigate whether the effectiveness of rhBMP-2 varied among patients who had anterior lumbar fusion or posterior lumbar fusion, found no evidence of a difference in the RRs for successful fusion (P=0.88) or any other outcome at 24 months across surgery types.	Return to work, hospital stay, operating timeadverse events,	There was no difference in duration of hospital stay (mean difference, -0.15 days [CI, -0.33 to 0.03 days]) or that rhBMP-2 surgery increased the probability of returning to work or usual activity earlier compared with ICBG (RR at 24 months, 1.01 [CI, 0.88 to 1.17]). Using rhBMP-2 shortened operating times by 21 minutes (CI, 15 to 27 minutes) from an average of 135 minutes.	-	-	Simmonds, Mark C.; Brown, Jennifer V. E.; Heirs, Morag K.; Higgins, Julian P. T.; Mannion, Richard J.; Rodgers, Mark A.; Stewart, Lesley A.. Safety and effectiveness of recombinant human bone morphogenetic protein-2 for spinal fusion: a meta-analysis of individual-participant data. Ann. Intern. Med.. 2013	-	-	Population: Degenerative disc disease. Mean age of participants in trials 40 - 70 years. Anterior lumbar fusion, posterior lumbar fusion, anteror cervical fusion. Comments: This is another large scale, independent meta-analysis based on Individual Participant data obtained from manufacturer and published data conducted as part of The Yale University Open Data Access (YODA) Project. The paper included sound methodology, including: search strategy, inclusion criteria, statistical analysis for heterogeneity, bias and quality of studies and presenetation of data as CI. The results of the metanalysis demonstrated that at 24 months, rhBMP-2 increases fusion rates, reduces pain by a clinically insignificant amount, and increases early postsurgical pain compared with ICBG. Evidence of increased cancer incidence is inconclusive . The results of the study are generalisable and main limitations of the meta-analysis, as noted by the authors, include lack of blinding for outcome assessment and ascertaintment of harms was poor in trials. Observational studies included in the reveiw were diverse and at risk of bias. Most studies included had links to industry.

2++	Systematic	508 (207 intervention vs 301 comparator group)	rhBMP2 in contained in one of the 3 different types of intradiscal implants 1. Threaded, tapered titanium cages (burkus 2002); 2. Threaded, cylindrical allograft bone dowels (Burkus 2005 and Burkus 2002) 3. Threaded, cylindrical titanium cages (Burkus 2004) undergoing ALIF fusion	Safety of the intervention	Retrograde ejaculation	Combining the data from the 5 trials, RE was reported in 7 (3.4%) of the 207 patients who received the rhBMP-2 treatment compared with 5 (1.7%) of the 301 patients who received the autograft or lumbar disc treatment (p = 0.242, Fisher exact test). Retrograde ejaculation occurred at the highest rates in the earliest clinical trial (Burkus et al 2002). Of the 146 men, 6 (4.1%) developed RE postoperatively. In subsequent studies, the rates of RE ranged from 0% to 2.1%. The difference in surgical approaches was significant (p = 0.007, Fisher exact test). Cases of RE were higher in group who underwent transperitoneal approach (5/58, 8.6%) compared to 1.6% (7/445) patients who underwent a retroperitoneal spinal exposure; There was no difference in the rate of RE based on the lumbar level exposed (p = 0.739). Multivariate analyses were consistent with the conclusions from Fisher exact tests.	0	-	-	-	Burkus, J. Kenneth; Dryer, Randall F.; Peloza, John H.. Retrograde ejaculation following single-level anterior lumbar surgery with or without recombinant human bone morphogenetic protein-2 in 5 randomized controlled trials: clinical article. J Neurosurg Spine. 2013	-	None mentioned	Population: Symptomatic degenerative disc disease at the L4–5 or L5–S1 levels suitable for a for a single-level stand-alone ALIF. 19 - 70 years. Subgroups by surgical approach-retroperitoneal vs transperitoneal. Comments: This is a study reporting on the complication of retrograde ejaculation after interbody surgery among groups treated with rhBMP-2, ICBG or lumbar arthroplasty implant. Results from 5 RCTs are combined and presented as rates of complications and difference between tested using Fischer exact test. Combining the data from the 5 trials, retrograde ejaculation (RE) was reported in 7 (3.4%) of the 207 patients who received the rhBMP-2 treatment compared with 5 (1.7%) of the 301 patients who received the autograft or lumbar disc treatment (p = 0.242, Fisher exact test). RE occurred at the highest rates in the earliest clinical trial (Burkus et al., 2002). Of the 146 men, 6 (4.1%) developed RE postoperatively. In subsequent studies, the rates of RE ranged from 0% to 2.1%. The difference in surgical approaches was significant (p = 0.007, Fisher exact test). Cases of RE were higher in group who underwent transperitoneal approach (5/58, 8.6%) compared to 1.6% (7/445) patients who underwent a retroperitoneal spinal exposure. There was no difference in the rate of RE based on the lumbar level exposed (p = 0.739). Multivariate analyses were consistent with the conclusions from Fisher exact tests. The generalisability of the results is limited for number of reasons, including: studies selected were not based on systematic search resulting in publication bias. The studies includes 5 RCTs conducted by the manufacturer. The quality of studies is not evaluated and there is no detail on the patient selection or how the heterogeneity was analysed suggesting possibility of bias. The author of the paper is the author of the 4 RCTs included in the paper and is consultant for the manufacturer and investigator in the studies included.
1+	Systematic	1342	rhBMP2	Clinical effectiveness of the intervention	Fusion failure, Reoperation,The Oswestry Disability Index The SF-36 physical component summary. The numeric rating scores for back pain. The numeric rating scores for leg pain. Adverse events.	Compared with ICBG, the use of rhBMP-2 significantly decreased the risk of fusion failure at all time intervals (6 months: p<0.0001, RR = 0.55, 95 % CI = 0.42–0.72; 12 months: p = 0.0003, RR =0.53, 95 % CI = 0.37–0.75; 24 months: p<0.00001, RR = 0.31, 95 % CI = 0.21–0.46) and the rate of reoperation (p = 0.0001, RR = 0.52, 95 % CI = 0.37–0.72). There was no statistical difference in clinical improvement on the Oswestry Disability Index, although a favorable trend in the rhBMP-2 group was found (p = 0.12, RR = 0.73, 95 % CI = 0.49–1.08). The result of subgroup analysis on reoperation demonstrated that, in both subgroups, the rate of reoperation in rhBMP-2 group was significantly lower than that of the ICBG group (anterior subgroup: p = 0.02, RR = 0.48, 95 % CI = 0.26–0.89; posterolateral subgroup: p = 0.002, RR = 0.54, 95 % CI = 0.36–0.79).. Due to different data formats, meta-analysis on adverse events was not performed.	None specifically mentioned in the study	-	-	-	Chen, Zhiguang; Ba, Gen; Shen, Tao; Fu, Qin. Recombinant human bone morphogenetic protein-2 versus autogenous iliac crest bone graft for lumbar fusion: a meta-analysis of ten randomized controlled trials. Arch Orthop Trauma Surg. 2012	Authors mention Retrograde ejacuation and others but details are not mentioned, The analysis was carried out due to differenr data formats	As in primary outcome measure	Population: Lumbar fusion for degenerative disc disease. Mean age of participants in trials 40 - 70 years. Instrumental PLF, PLIF, posterolateral interbody fusion (IF) and ALIF. Comments: This is a independently conducted, meta-analysis of studies comparing rhBMP-2 with ICBG published up to Feb 2013. The study has clear eligibility criteria with inclusion and exclusion criteria, but no primary and secondary outcomes. The risk of bias assessment and evaluation of validity was done using the 12 criteria recommended by the Cochrane review group and clinical relevance was assessed using questions recommended by Cochrane Back Review Group. Robust statistical methods were used to assess heterogeneity. A random effect model was used in the meta-analysis and funnel plots and statistics tests (Egger's and Begg's) were done to assess publication bias. Compared with ICBG, the use of rhBMP-2 significantly decreased the risk of fusion failure at all time intervals (6 months: p<0.0001, RR = 0.55, 95 % CI = 0.42–0.72; 12 months: p = 0.0003, RR =0.53, 95 % CI = 0.37–0.75; 24 months: p<0.00001, RR = 0.31, 95 % CI = 0.21–0.46) and the rate of reoperation (p = 0.0001, RR = 0.52, 95 % CI = 0.37–0.72). There was no statistical difference in clinical improvement on the ODI, although a favorable trend in the rhBMP-2 group was found (p = 0.12, RR = 0.73, 95 % CI = 0.49–1.08). Subgroup analyses stratified by the type of surgical procedure yielded similar results. Due to different data formats, meta-analysis on adverse events was not performed. However, some limitation of this review include: low quality of studies included as evaluated using Cochrane review group quality assessment tool, three of ten included studies were rated to be with ''high risk of bias", lack of double blinding and the majority of the trials were sponsored by manufacturers of BMP.

2+	Systematic	780	rhBMP-2	Clinical effectiveness of the intervention	Complications from rhBMP2 compared with ICBG	rhBMP-2 use in spine fusion ranging from 10% to 50% depending on approach. Anterior cervical fusion with rhBMP-2 has an estimated 40% greater risk of adverse events with rhBMP-2 in the early postoperative period, including life-threatening events. After anterior interbody lumbar fusion rates of implant displacement, subsidence, infection, urogenital events, and retrograde ejaculation were higher after using rhBMP-2 than controls. Posterior lumbar interbody fusion use was associated with radiculitis, ectopic bone formation, osteolysis, and poorer global outcomes. In posterolateral fusions, the risk of adverse effects associated with rhBMP-2 use was equivalent to or greater than that of iliac crest bone graft harvesting, and 15% to 20% of subjects reported early back pain and leg pain adverse events; higher doses of rhBMP-2 were also associated with a greater apparent risk of new malignancy.	None	-	-	-	Carragee, Eugene J.; Hurwitz, Eric L.; Weiner, Bradley K.. A critical review of recombinant human bone morphogenetic protein-2 trials in spinal surgery: emerging safety concerns and lessons learned. Spine J. 2011	As in primary outcome measure	As in primary outcome measure	Population: Patients with lumbar and cervical spondylosis. PLF, PLIF, ALIF and anterior cervical fusion. Comments: This is a systematic review of studies evaluating safety of rhBMP-2 compared with ICBG. This review has clear onbjectives, serach methods and statistical methods. However, the methodology doesn't include polling of data to estimate the effect size and regression methodology to adjust for commforunders. The results show that, compared to previously reported, low levels of side effects of rhBMP-2, authors found that rhBMP-2 was associated with high levels of adverse effects ranging from 10% to 50% depending on approach. Anterior cervical fusion with rhBMP-2 has an estimated 40% greater risk of adverse events with rhBMP-2 in the early postoperative period, including life-threatening events. The generalisation of study are limited in that most of the studies included in the review are retrospective case series with poor patient selection methods and lack of blinding in patient reported measures. However, findings from this study are significant as it highlighted the side effects with rhBMP-2
1+	Systematic	184324 (28,815 experimental, 155,509 controls)	rhBMP-2	Safety of the intervention	General complications, heterotopic ossification (HO), retrograde ejaculation, cervical swelling, and cancer rates with the use of rhBMP-2 in lumbar and cervical spine fusion.	There was a significantly higher risk of general complications with rhBMP-2 compared to iliac crest bone graft (ICBG) with an odds ratio (OR) of 1.78 (95 %CI 1.20–2.63), (p = 0.004). The odds ratio for HO was 5.57 (95 %CI 1.90–16.36), (p = 0.002), for retrograde ejaculation 3.31 (95 %CI 1.20–9.09), (p = 0.020), and for cervical swelling 4.72 (95 %CI 1.42–15.67), (p = 0.011), all significantly higher in the rhBMP-2 group. The pooled odds ratio for new onset of tumor was 1.35 (95 %CI 0.93–1.96), which represents no statistically significant difference between the groups (p = 0.111).	None	-	-	-	Vavken, Julia; Mameghani, Alexander; Vavken, Patrick; Schaeren, Stefan. Complications and cancer rates in spine fusion with recombinant human bone morphogenetic protein-2 (rhBMP-2). Eur Spine J. 2015	-	-	Population: Not specified. Mean age was 51.1 +/- 1.8 years. Comments: This is an independently conducted, well designed and presented meta-analysis of studies comparing complications of rh BMP-2 with ICBG using PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) and QUOROM (Quality Of Reporting Of Meta-analysis) statements. The study has clear eligibility criteria with inclusion and exclusion criteria, primary and secondary outcomes. Internal validity of each study was evaluated with modified Jadad scale and publication bias was assessed using funnel plot and Egger's weighted regression technique. There was a significantly higher risk of general complications with rhBMP-2 compared to iliac crest bone graft (ICBG) with an odds ratio (OR) of 1.78 (95 %CI 1.20–2.63), (p = 0.004). The OR for HO was 5.57 (95 %CI 1.90–16.36), (p = 0.002), for retrograde ejaculation 3.31 (95 %CI 1.20–9.09), (p = 0.020), and for cervical swelling 4.72 (95 %CI 1.42–15.67), (p = 0.011), all significantly higher in the rhBMP-2 group. The pooled OR for new onset of tumour was 1.35 (95 %CI 0.93–1.96), which represents no statistically significant difference between the groups (p = 0.111). However, some limitations of this review, as recognised by the authors, include low quality of studies, lack of double blinding and the majority of trails were sponsored by the manufacturers.

2+	Case series	33 patients receiving rhBMP-2 in addition to either local bone autograft or ICBG (rhBMP-2 cohort) and 42 patients receiving only local bone autograft or ICBG (control cohort) for one or two-level dorsal lumbar fusion	rhBMP-2 alone or in addition to either local bone autograft or ICBG for for one or two-level dorsal lumbar fusion includingpost erolateral fusion (PLF), posterior lumbar interbody fusion (PLIF), and transforamin al lumbar interbody fusion (TLIF	Cost effectiveness	Postoperative 1-year cost/utility ratios and the incremental cost effectiveness ratio (ICER) to assess for cost effectiveness using a threshold of $100,000/QAL Y gained	1-year cost-utility ratio (Total Cost/ΔQALY) for the control cohort was significantly lower ($143,251/QALY gained) than that of the rhBMP-2 cohort ($272,414/QALY gained) (P<0.01). At 1-year follow-up, the control group dominated the ICER compared to the rhBMP-2 group.	1-year postoperative Health outcomes assessed based on Visual Analogue Scale (VAS), Pain Disability Questionnaire (PDQ), Patient Health Questionnaire (PHQ-9), and EuroQol-5 Dimensions (EQ-5D) questionnaires. Direct medical costs estimated using. Medicare national payment amounts and indirect costs based on patient missed work days and patient income	No details available from the abstarct	-	-	Alvin, Matthew D.; Derakhshan, Adeeb; Lubelski, Daniel; Abdullah, Kalil G.; Whitmore, Robert G.; Benzel, Edward C.; Mroz, Thomas E.. Cost-utility Analysis of One and Two-level Dorsal Lumbar Fusions With and Without Recombinant Human Bone Morphogenic Protein-2 at 1-year Follow-up. J Spinal Disord Tech. 2014	-	-	Population: Not specified. Comments: A smalll study of 75 patients in which 33 receivied rhBMP-2 in addition to either local bone autograft or ICBG (rhBMP-2 cohort) and 42 patients receiving only local bone autograft or ICBG (control cohort) for one or two-level dorsal lumbar fusion. Study is presented with good methodolgy with description tools used in measuring health outcomes and estimation of various costs. Health outcomes were assessed based on Visual Analogue Scale (VAS), Pain Disability Questionnaire (PDQ), Patient Health Questionnaire (PHQ-9), and EuroQol-5 Dimensions (EQ-5D) questionnaires. Direct medical costs were estimated using Medicare national payment amounts and indirect costs were based on patient missed work days and patient income. Postoperative 1-year cost/utility ratios and the incremental cost effectiveness ratio (ICER) were calculated to assess for cost effectiveness using a threshold of $100,000/QALY gained.The 1-year cost-utility ratio (Total Cost/ΔQALY) for the control cohort was significantly lower ($143,251/QALY gained) than that of the rhBMP-2 cohort ($272,414/QALY gained) (P<0.01). At 1-year follow-up, the control group dominated the ICER compared to the rhBMP-2 group. By 2 years, the control cohort was considered cost effective ($71,625/QALY gained) compared to ($136,207/QALY) f or the rhBMP-2 cohort. Main limitations of the study include: it was not a direct comparision of rhBMP-2 against ICBG and lack of randomisation in patient selection methods and lack of blinding, especially for the patient reported measures. It is not clear form the methods if there was time horizon discounting for estimation of cost effectiveness. .

3	Case series	57 patients (Forty-eight patients (84.2%) had undergone previous cervical surgery, and 42.1% had a preexisting nonunion).	rhBMP-2	Clinical effectiveness of the intervention	Fusion status was determined by bony bridging on computed tomography scans, absence of radiolucency around instrumentation, and absence of motion on lateral flexion/extension radiographs	Postoperatively, 51 patients (89.5%) developed fusion, while 6 (10.5%) had radiographic evidence of pseudarthrosis	The neurologic symptoms of Visual analog scale (VAS) pain and Neck Disability Index (NDI)	The neurologic symptoms of 50 patients (87.7%) resolved postoperatively, and Visual analog scale (VAS) pain and Neck Disability Index (NDI) 26 scores improved significantly from baseline. VAS (0 = no pain, 10 = maximum pain) fell from a mean of 6.8 ± 1.8 to 3.9 ± 3.1 (P < 0.001) and NDI improved from a mean of 50.6 ± 20.6 to 36.7 ± 19.3 (P = 0.002).	-	-	Dorward, Ian G.; Buchowski, Jacob M.; Stoker, Geoffrey E.; Zebala, Lukas P.. Posterior Cervical Fusion with Recombinant Human Bone Morphogenetic Protein-2: Complications and Fusion Rate at Minimum Two-Year Follow-Up. J Spinal Disord Tech. 2013	-	-	Population: Common surgical indications were adjacent-level spondylosis following fusion (35 cases, 61.4%) and symptomatic pseudarthrosis (22 cases, 38.6%). Other indications included basilar invagination, postsurgical kyphosis, rheumatoid arthritis, and chronic fracture non-union. Mean age 56.1 ± 13.6 years. Comments: This is a retrospective study of 57 patients who underwent cervical fusion using rhBMP-2. Forty-eight patients (84.2%) had undergone previous cervical surgery, and 42.1% had a pre-existing non-union).There was no comparator group and the patient selection method is poorly described. The results show that 87% achieved fusion and there was significant improvement in VAS and NDI score from baseline.The primary outcomes are not anlaysed according to previous surgery although the majority of the patients had undergone a cervical fusion surgery. The generalisability of the results is limited by lack of comparator, lack of blinding, and small sample size.

1-	Systematic	0	RhBMP-2 (RhBMP),Iliac crest bone graft (ICBG),,local bone alone (LBG), demineralized bone matrix with local bone (DBM), local bone with corticocancellous allograft chips (CCA)	Cost effectiveness	cost/QALY	Incremental cost-effective ratio for each graft option when compared with living with chronic back pain was $21,308/QALY for ICBG, $16,595/QALY for RhBMP, $21,204/QALY for LBG, $21,287/QALY for DBM, and $28,153/QALY for CCA. However the sensitivity analysis showed RhBMP was not the most cost-effective option if the revision rate is significantly raised. and if the cost of treatment with RhBMP rises >$42,250. In that case LBG becomes the likely cost-effective treatment.	-	-	-	-	Virk, Sohrab; Sandhu, Harvinder S.; Khan, Safdar N.. Cost effectiveness analysis of graft options in spinal fusion surgery using a Markov model. J Spinal Disord Tech. 2012	-	As in primary outcome measure	Population: Lumbar degenerative spondylolisthesis. Comments: This is a good quality study evaluating cost effectiveness of grafts used in lumbar surgery, including: rhBMP-2, ICBG, local bone alone (LBG), demineralised bone matrix with local bone (DBM) and local bone with corticocancellous allograft chips (CCA). Base comparison was chronic back pain and an incremental cost effectiveness ratio (ICER) was calculated for each graft option. ICER is calculated as (cost fusion-cost chronic back pain)/(QALY fusion-QALY chronic back pain). Sensitivity analysis was performed to stress the inputs within the Markov model. Specifically, the revision rates, costs associated with primary fusion surgery and QALY values for each graft option were varied. One-way sensitivity analysis for each of these 3 variables was performed. Two-way sensitivity analysis was also performed to evaluate those variables that seemed to most affect the results of the model. All costs are appropriately adjusted as 2011 dollars. For rhBMP-2 and ICBG, the rates of fusion, revision rates, and complications are based on studies by Cahill et al. (2011), Dimar et al. (2009), Vaccaro et al. (2007). Result show that cost-effective ratio when compared with living with chronic back pain was $16,595/QALY for rhBMP-2- the most cost-effective graft option. However the sensitivity analysis shows that rhBMP-2 is not the most cost-effective option if the revision rate is significantly raised. If the cost of treatment with rhBMP-2 rises >$42,250, then LBG becomes the likely cost effective treatment. This is significant considering the findings from a recent population level study by Savage et al. (2015) which showed that the 90 day complications rate in group using rhBMP-2 for lumbar spinal fusion was significantly higher than group using non-BMP methods (RR1.108 (CI 1.060–1.158). Limitations of this study include the lack of long-term follow up of revision rates and the number of patients involved in evaluating QALY data points was often small.

0	Systematic	Not mentioned	rhBMP2	clinical effectiveness of the intervention	Bone fusion and clinical outcome measures as measured using ODI , back pain and leg pain scores	Anterior lumbar fusions: rhBMP2 likely associated with an increased rate of radiographic arthrodesis when compared with ICBG. However, this does not necessarily translate to an improvement in clinical outcomes. There are potential complications that are specific to rhBMP-2 utilization including osteolysis and retrograde ejaculation. rhBMP-2 may also be associated with lumbar plexopathy when utilized in the transpsoas lumbar fusion cases. Single level posterior lumbar fusions: rhBMP-2 may improve the rates of radiographic arthrodesis in posterior lumbar fusion procedures, particularly in patients at high-risk for pseudarthrosis such as smokers but is based one study by Glassman et al. RhBMp2 specific complications include increased early back and leg pain,ectopic bone formation and radiculitis. Multiple level posterior lumbar fusions: There are very limited studies which examine only multilevel posterior degenerative lumbar fusions. The existing data provides evidence that rate of complications rhBMP-2 is similar to ICBG. Cervical fusions: There are relatively few high quality studies assessing the risks and benefits of rhBMP-2 utilisation in anterior cervical spine fusion surgery. The use of rhBMP-2 in posterior cervical fusions is likely safer than in anterior approaches. Anterior cervical fusion is associated with higher rate of complications including dysphagia, dysphonia and wound infection.	none	-	-	-	Walker, Brett; Koerner, John; Sankarayanaryanan, Sriram; Radcliff, Kris. A consensus statement regarding the utilization of BMP in spine surgery. Curr Rev Musculoskelet Med. 2014	-	as in primary outcome	Population: Spinal fusion including lumbar and cervical. High risk groups- cancer, smokers, osteomyelitis. Comments: This is an overview of studies evaluating rhBMP-2. The study doesn't include: a search strategy, patient selection, evaluation of study quality or statistical methods to pool data. Hense generalisability of this study is limited.

1+	Systematic	496 pateients with eithet rhBMP2 or ICBG	rhBMP2 or ICBG	Clinical effectiveness of the intervention	radiological fusion	Of the 496 patients with clinical and radiographic data at 1 and 2 year follow-ups were identified. Of these, 5.5% [95%CI: 3.7; 8.3] had radiographic nonunion which did not require reoperation	Clinical outcomes (Oswestry Disability Index (ODI); Numeric Rating Scales (NRS) for back and leg pain	Patients with fusion had better improvements in ODI (P<0.001) and NRS back pain scores (P<0.001). The overall percentage of fused patients with ODI and NRS back pain scores that exceeded the criteria for minimal clinically important differences (MCID) was also significantly higher than that of patients with nonunion (ODI, OR = 2.7, P = 0.019; NRS back pain, OR = 3.5, P = 0.033). However, the predictive values of fusion for clinical outcomes were poor, with low specificity and low negative predictive values.	-	-	Noshchenko, Andriy; Lindley, Emily M.; Burger, Evalina L.; Cain, Christopher M. J.; Patel, Vikas V.. What is the Clinical Relevance of Radiographic Nonunion after Single-level Lumbar Interbody Arthrodesis in Degenerative Disc Disease? A Meta-Analysis of the YODA Project Database. Spine. 2015	None mentioned/included	as in primary outcome	Population: Patients with lumbar arthrodesis, using rhBMP-2 or ICBG. Comments: This is a good quality meta-analysis of RCTs comparing ICBG and rhBMP-2 to enaluate relation between radiological bone fusion, which is used as primary outcome measure, and clinical outcomes. The results show that patients who had radiological fusion had significantly better patient outcomes, including clinical outcomes (Oswestry Disability Index (ODI); Numeric Rating Scales (NRS) for back and leg pain). However, the abstact doesnt present the results by ICBG and rhBMP-2. The study is limited in that patients in the RCTS were not blinded, so role bias cannot be ruled out. Authors also note that the predictive values of fusion for clinical outcomes were poor, with low specificity and low negative predictive values and need for direct evaluation of patient related outcomes in future studies.

3	Case series	1997-2006	rhBMP2 for long fusions to the sacrum	Clinical effectiveness of the intervention	Oswestry Disability Index and 3 domains of the Scoliosis Research Society score . BMP dose	The fusion rates for BMP group were 93.5% and 71.9% for the ICBG group. The rate pseudarthrosis was 6.4% (2/31) in the BMP and 28.1% (9/32) in the ICBG group (P = 0.04). Oswestry Disability Indexes were similar between groups. However, the BMP group demonstrated superior sum composite Scoliosis Research Society scores in pain, self-image and function domains (P = 0.02). The concentration and dosage of recombinant human bone morphogenetic protein 2 (rhBMP-2) used seems to have an effect on the rate of fusion and pseudarthrosis rate because no patient receiving more than 5 mg per level had apparent or detected pseudarthroses (n = 20/20).	None	-	-	-	Kim, Han Jo; Buchowski, Jacob M.; Zebala, Lukas P.; Dickson, Douglas D.; Koester, Linda; Bridwell, Keith H.. RhBMP-2 is superior to iliac crest bone graft for long fusions to the sacrum in adult spinal deformity: 4- to 14-year follow-up. Spine. 2013	-	As in primary outcome measure	Population: Ambulators who were candidates for long fusions (thoracic as the upper level) to the sacrum. Adults. Comments: A retrospective case series of 63 patients treated for long fusion to sacrum using rhBMP-2 or ICBG. The rhBMP-2 group had higher rate of radiological fusion but ODI scores were simialr across both groups, suggesting that radiological fusion didn't translate into better patient outcomes. The authors conclude that patients receiving higher dose of rhBMP-2 had better success but there is lack of infomation on baseline characteristics of high dose and also the clinical reasons for high dose. Overall, the generalisability of study results is limited due to patient selection, lack of bliniding and lack of robust statistical methods in estimating differences in outcomes between the groups.
2-	Case series	148	rhBMP2 for single-level lumbar fusion.	Clinical effectiveness of the intervention	Radiologically estimated Solid fusion and clinical outcome measures including Oswestry Disability Index, SF-36, back, and leg pain scores	At 2 years postoperatively, solid fusion was demonstrated in all 55 nonsmokers in the rhBMP-2 group (100%). Successful fusion was seen in 20 of 21 smokers in the rhBMP-2 group (95.2%). Fusion was achieved in 48 of 51 nonsmokers in the iliac crest bone graft (ICBG) group (94.1%), but only 16 of 21 smokers (76.2%) in the ICBG group.	none	-	-	-	Glassman, Steven D.; Dimar, John R.; Burkus, Kenneth; Hardacker, James W.; Pryor, Philip W.; Boden, Scott D.; Carreon, Leah Y.. The efficacy of rhBMP-2 for posterolateral lumbar fusion in smokers. Spine. 2007	-	-	Population: Lumbar Degenerative disc disease . Adults. Smokers. Comments: A retrospective study comapring outcome of lumbar fusion using rhBMP-2 in patients with smoking and non-smokers. At 2 years postoperatively, solid fusion was demonstrated in all 55 non-smokers in the rhBMP-2 group (100%). Successful fusion was seen in 20 of 21 smokers in the rhBMP-2 group (95.2%). Fusion was achieved in 48 of 51 nonsmokers in the iliac crest bone graft (ICBG) group (94.1%), but only 16 of 21 smokers (76.2%) in the ICBG group. The generalisability of study is limited due to retrospective patient selection methods , with no randomisation and lack of explanation for excluding patients who had similar surgery but were excluded from the study.

2+	Case series	Sixty-two patients. .	rhBMP-2 on an absorbable collagen sponge.	0	Radiological bone fusion and clinical outcomes. For the radiologic assessment of fusion rate, 3 criteria were used: the presence of trabeculated bone between transverse processes, no implant loosening, and less than 2° of movement on lateral flexion and extension films. Clinical outcomes were assessed using a VAS (0 no pain, 10 worst possible pain) for back and leg pain before surgery and at 6-week, 6-, 12-, and 24-month follow-up	Patients in rh-BMP-2 and autogenous bone graft group achieved a 100% fusion rate for both single- and multilevel revision PLF. In BMAA group the fusion rate for single level revision PLF was 100% (7/7) and the multilevel fusion rate of 64% (7/11). There was a significant decrease between preoperative and 2-year postoperative VAS scores in all groups (P 0.001), but no significant difference among groups at all time points.	Complicationrate	8% (2/24) of rh BMP2, 22% of BAAA (4/18) and 10% (2/20) of autograft groups required additional surgical intervention and four patients (20%) in autograft group complained of persistent donor-site pain at 2-year follow-up	-	-	Taghavi CE1, Lee KB, Keorochana G, Tzeng ST, Yoo JH, Wang JC.. Bone Morphogenetic Protein-2 and Bone Marrow Aspirate With Allograft as Alternatives to Autograf tin Instrumented Revision Posterolateral Lumbar Spinal Fusion . Spine. 2010	8% (2/24) of rh BMP2, 22% of BAAA (4/18) and 10% (2/20) of autograft groups required additional surgical intervention and four patients (20%) in autograft group complained of persistent donor-site pain at 2-year follow-up	As in primary outcome results	Population: Indications for revision surgery included symptomatic pseudarthrosis (pain and/or instability) following a previous PLF for degenerative conditions of the lumbar spine, such as degenerative disc disease, stenosis, or spondylolisthesis. Group 1 contained 24 patients (13 single- [group 1A] and 11 multilevel [group 1B]) who underwent instrumented revision PLF using rhBMP-2 on an absorbable collagen sponge. Group 2 included 18 patients (7 single- [group 2A] and 11 multilevel [group 2B]) with procedures using bone marrow aspirates in conjunction with allograft (BMAA). Group 3 consisted of 20 patients (10 single- [group 3A] and 10 multilevel [group 3B]) with procedures using autograft. Comments: This is a small restrospective case reveiw of 62 consecutive patients who underwent instrumented revision posterolateral spine fusion between January 2002 and December 2006. Patients were treated with either rhBMP-2 on an absorbable collagen sponge (n-24), Bone marrow aspirates used in conjunction with allograft(BMAA) (n=18) or autograft (n=20). The exact source of autograft bone for Group 3 was not clearly defined. All 3 cohorts received supplemental local bone, static and dynamic radiographs were used to assess fusion and were reviewed by 3 blinded independent reviewers with a diagnosis of non-union based on either surgical exploration if revision performed or radiographic findings. Clinical outcome was determined through VAS scores. The baseline characteristics of the three groups were similar in demography and risk factors. The results show that bone fusion rate in rhBMP-2 and autograft was 100% for both single and multi-level fusion. The BAAA had lower fusion rate for multi-level fusion but 100% for single level fusion. There was no difference betwen groups for clinical outcomes. Higher proportion of patients in BAAA had a revision surgery and (20% in autograft group complained of persistent donor-site pain at 2-year follow-up.

2+	Case series	One hundred ninety-five patients were divided into 4 groups depending on fusion material and the presence/ absence of fusion-related risk factors for nonunions;	rhBMP2	clinical effectiveness	Fusion rate and rate of fusion	The time to fusion was significantly faster in group B than in group D in patients with no history of smoking (P<0.05), hypertension (P<0.01), or other significant comorbidity (P<0.05). The time to complete fusion was also significantly faster in group B than in group D in patients under the age of 65 (P<0.05), patients undergoing primary surgery (P<0.05), single-level surgery (P<0.01), no smoking history (P<0.05), no diabetes mellitus (P<0.01), no hypertension (P=0.001), no osteoporosis (P<0.01), and no significant comorbidity (P<0.01). Although the fusion rate was higher in group B than in group D, with the exception of sex and single-level surgery, there were no significant differences between groups B and D. Although initial fusion mass and time to solid fusion was faster in group A than in group C, there were no significant differences between groups A and C. In addition, fusion rates were higher in group C than in group A, looking at all factors except revision surgery, but the differences were not statistically significant.	None	-	-	-	Kwang-Bok Lee, MD, PhD,*w Jared S. Johnson, MD,* Kyung-Jin Song, MD,w Cyrus E. Taghavi, BS,* and Jeffrey C. Wang, MD*. Use of Autogenous Bone Graft Compared With RhBMP in High-risk Patients A Comparison of Fusion Rates and Time to Fusion. J Spinal Disord Tech V. 2013	None included	As in primary outcome measure	Population: Degenerative lumbar spine diseases. Adults. Group A defined as rhBMP-2 used in the presence of high-risk factors (FRRF), group B was defined as rhBMP-2 used in the absence of FRRF, group C was defined as autograft used in the presence of FRRF, and group D was defined as autograft used in the absence of FRRF. Comments: A retrospective study comapring rhBMP-2 and autograft in patients with risk factors for non-union. At 24 month follow up, although time to fusion was faster in group A than in group C in all fusion-related risk factors (age, sex, revision, fusion level, smoking, DM, osteoporosis, and comorbidity), there was no statistically significant difference between groups A and C. Similarly, fusion rate was higher in group A than in group C in other fusion related risk factors, except revision surgery but there was no statistically significant difference between groups A and C in all fusion-related risk factors. The generalisability of the study is limited due to retrospective patient selection methods, with no randomisation and lack of explanation of patiants who had simialr surgery but were excluded from the study.
2-	Case series	127	rhBMP-2 and allograft	clinical effectiveness	Fusion rate and rate of fusion, clinical outcome, VAS, perioperative complications and revision rate	At the end of 24 month follow up period the fusion rate and fusion time were similar in groups A and C; however, these were lower than that observed in group B. Clinical outcomes were similar amongst the groups, there were no significant differences in VAS and perioperative complication rate between groups A and C. Similarly there was also no difference in fusion rate between the group A and group C when analysed by gender, comorbidity, osteoporosis, previous surgery for fusion, and smoking	None	-	-	-	Kwang-Bok Lee, Cyrus E. Taghavi, Margaret S. Hsu, Kyung-Jin Song, Jeong Hyun Yoo. Gun Keorochana, Stephanie S. Ngo, Jeffrey C. Wang . The efficacy of rhBMP-2 versus autograft for posterolateral lumbar spine fusion in elderly patients. Eur Spine J . 2010	-	As in primary outcome measure	Population: Degenerative lumbar spine diseases. Adults. Subjects in group A (n = 34) consisted of patients 65 + years who received rhBMP-2 and allograft. Group B (n = 52) was composed of patients under 65 years of age with rhBMP-2 and allograft. Subjects in group C (n = 41) were 65 + years with autograft use. Comments: A retrospective study comparing rhBMP-2 with allograft and autograft in older patients. At the end of 24 month follow up period the fusion rate and fusion time were similar in groups A and C; however, these were lower than that observed in group B. Clinical outcomes were similar amongst the groups, there were no significant differences in VAS and perioperative complication rate between groups A and C. Similarly, there was also no difference in fusion rate between the group A and group C when analysed by gender, comorbidity, osteoporosis, previous surgery for fusion and smoking. The generalisability of the study is limited due to retrospective patient selection methods , with no randomisation and lack of explanation of patiants who had simialr surgery but were excluded from the study.

2-	Retrospective Cohort	460773	BMP for spinal fusion	Clinical effectiveness of the intervention	Complications were categorized into 1 of 10 groups: respiratory; peripheral vascular; CNS; hematoma; accidental cut, puncture, or hemorrhage during the procedure; complications of the operative wound, including infection; other; CSF leak; deep vein thrombosis; and mechanical complication of an implant or graft In addition, we determined the average Charlson Comorbidity Index (CCI) for patients without complications and for patients with 1 or more complications.	The overall complication rate in the BMP group was 18.2% compared with 18.7% in the control group. The RR of BMP use compared with no BMP use was 0.976 (CI 0.963–0.989) (p < 0.001). Age: In both treatment groups, patients older than 65 years had a significantly higher rate of postoperative complications than the younger patients (p < 0.001) . In patients younger than 65 years, the RR of developing a complication with the use of BMP was 1.042 (CI 1.017–1.067), whereas in the patients ≥ 65 years old, the opposite was true (RR 0.950 [CI 0.935–0.065]. Effect of Sex: For both males and females, the complication rates were lower in the BMP group than in the control group. The RRs of BMP use compared with no BMP use were 0.974 (CI 0.953–0.995) in males and 0.976 (CI 0.960–0.993) in females. The RR was significantly lower in the BMP group for females but not for males (p < 0.001). Charlson Comorbidity Index: Patients with 1 or more complications had a higher CCI than patients in all subgroups who had no complications. In addition, patients who underwent spinal fusion without BMP had a higher CCI than the patients in all subgroups who received BMP, except for female patients without a complication. Reoperations Rates: The overall 90-day reoperation rates were 1.84% in the control group and 2.03% in the BMP group . The RR of reoperation was 1.108 (CI 1.060–1.158), which was significant and indicates a lower rate in the control group. In both the control and BMP groups, patients younger than 65 years were more likely to have a reoperation than patients older than 65 years (p < 0.001).	None	-	-	-	Jason W. Savage, MD ,1 Mick P. Kelly, MD ,2 Scott A. Ellison, MB A,3 and Paul A. Anderson, MD 2. A population-based review of bone morphogenetic protein: associated complication and reoperation rates after lumbar spinal fusion. Neurosurg Focus . 2015	As in primary outcome measure	As in primary outcome measure	Population: Patients who underwent lumbar fusion either using BMP or non BMP; all ages; subgroups by age and sex. Comments: This is an industry sponsored retrospective case series comparing complication rates in patients undergoing lumbar fusion using BMP vs non BMP. The objective and defintion of primary outcimes are well defined. Although patient selection method for patient using BMP is defined there is lack of clarity in patient selection methods of patients using non BMP. Also there is information on patients by different types of lumber surgery ie. anterior, lateral or posterior. The results show that the complication rates were significantly lower in BMP group. Authors also report complication in <65 yrs age group was higher in BMP group than non -BMP group but opposite was true in >65 yr age group. Also complication was significantly lower in the BMP group for females but not for males (p < 0.001). However for the reoperations rates: the overall 90-day reoperation rates were 1.84% in the control group and 2.03% in the BMP group . The RR of reoperation was 1.108 (CI 1.060–1.158), indicating a significantly lower rate in the control group. The results of the study are generalisable in the context of large sample size however as the study lacks info by types of lumbar surgery and also by types of non BMP it is not possible to conclude the safety of BMP by type of surgery. Also the data for the study patients is based on industry database the completeness of which is unknown and also patient selection methods are insufficent to make a recommendation.

Appendix Two

Literature search terms

Assumptions / limits applied to search:	
Original search terms:	Meta-analyses Systematic reviews Randomised controlled trials Prospective non-randomised clinical studies or other clinical studies Health economics studies
Updated search terms - Population	Anterior lumbar interbody fusion* Lumbar fusion* Spine fusion* Spinal fusion* ALIF Trans-foraminal lumbar interbody fusion* Transforaminal lumbar interbody fusion* Posterior Lumbar Interbody Fusion* Cervical fusion* TLIF PLIF Thoracic fusion* An additional search for the intervention, comparator and following population terms only is as follows: Pseudarthrosis Non-union* Non union* Nonunion* Revision* Revised Ununited fracture* Fail*
Updated search terms - Intervention	Bone morphogenetic protein* Bone morphogenetic protein-2 Inductos Infuse Dibotermin alfa BMP BMP-2 rhBMP-2 rBMP rhBMP

Updated search terms - Comparator	Iliac graft* Autologous graft* Autogenous graft* ABG Iliac bone Iliac crest
Updated search terms - Outcome	Not applicable
Inclusion criteria	**General inclusion criteria**
	In order of decreasing priority, articles will be selected based on the following criteria. 1.All relevant systematic reviews and meta-analysis in the last 5 years and those in 5-10 years period which are still relevant (e.g. no further updated systematic review available) 2.All relevant RCTs and those in the 5-10 years period which are still relevant (e.g. not superseded by a next phase of the trial/ the RCT is one of the few or only high quality clinical trials available) >>>> If studies included reaches 30, inclusion stops here 3.All relevant case control and cohort studies, that qualify after exclusion criteria >>>> If studies included reaches 30, inclusion stops here 4.All relevant non analytical studies (case series/ reports etc.) that qualify after exclusion criteria >>>> If studies included reaches 30, inclusion stops here
	Specific inclusion criteria
	Following articles were included per suggestion of the evidence reviewer: 1. Glassman, Steven D.; Dimar, John R.; Burkus, Kenneth; Hardacker, James W.; Pryor, Philip W.; Boden, Scott D.; Carreon, Leah Y.. The efficacy of rhBMP-2 for posterolateral lumbar fusion in smokers. Spine. 2007 2. Taghavi CE1, Lee KB, Keorochana G, Tzeng ST, Yoo JH, Wang JC.. Bone Morphogenetic Protein-2 and Bone Marrow Aspirate With Allograft as Alternatives to Autograf tin Instrumented Revision Posterolateral Lumbar Spinal Fusion . Spine. 2010 3. Kwang-Bok Lee, MD, PhD,*w Jared S. Johnson, MD,* Kyung-Jin Song, MD, Cyrus E. Taghavi, BS,* and Jeffrey C. Wang, MD*. Use of Autogenous Bone Graft Compared With RhBMP in High-risk Patients A Comparison of Fusion Rates and Time to Fusion. J Spinal Disord Tech V. 2013 4. Kwang-Bok Lee, Cyrus E. Taghavi, Margaret S. Hsu, Kyung-Jin Song, Jeong Hyun Yoo. Gun Keorochana, Stephanie S. Ngo, Jeffrey C. Wang . The efficacy of rhBMP-2 versus autograft for posterolateral lumbar spine fusion in elderly patients. Eur Spine J . 2010 5. Jason W. Savage, MD ,1 Mick P. Kelly, MD ,2 Scott A. Ellison, MB A,3 and Paul A. Anderson, MD 2. A population-based review of bone morphogenetic protein: associated complication and reoperation rates after lumbar spinal fusion. Neurosurg Focus . 2015
Exclusion criteria	**General exclusion criteria**
	Studies with the following characteristics will be excluded: 1. Does not answer a PICO research question 2. Comparator differs from the PICO 3. No relevant outcomes 4. Incorrect study type 5. Inclusion of outcomes for only one surgeon/doctor or only one clinical site (where studies with > one surgeon/doctor or one clinical site exist) 6. Narrative / non-systematic reviews (relevant referenced studies to be included)
	Specific exclusion criteria
	-