Clinical Effectiveness of Bone Morphogenetic Protein-2 (BMP-2)for Lumbar interbody Fusion: Meta-analysis of Controlled Studies

Kimberly Garrison
Simon Donell
Fujian Song*

Norwich Medical School,
Faculty of Medicine and Health Science,
University of East Anglia, Norwich, NR4 7TJ, UK.

May 2011

Acknowledgement: The work was supported by Medtronic, the manufacturer of BMP-2 products.

Abstract

Background

Bone morphogenetic protein (BMP) may be used to replace the use of autologous bone graft in lumbar interbody fusion. This paper aims to provide an updated meta-analysis of research evidence on the efficacy of safety of the use of BMP-2 for lumbar interbody fusion.

Methods

We systematically searched different databases to identify randomised (RCT) and non-randomised (non-RCT) controlled studies that compared BMP-2 and autograft (ABG) for lumbar interbody fusion. Risk of bias in primary studies and heterogeneity in meta-analyses were assessed. Random-effects model is used by default to pool results from individual studies.

Results

Five RCTs and seven non-RCT studies were included. The use of BMP-2 was associated with a higher rate of success fusion than the use of ABG (82.6% vs. 70.1%; $P<0.02$). There was no significant difference in the improvement in Oswestry Index score (pooled mean difference ($P=0.18$). The use of BMP-2 was associated with fewer secondary revisions (10.4% vs. 17.5%, $P=0.04$), shorter operating times ($P=0.003$) and hospital stay lengths ($P=0.02$). Pre-specified subgroup analyses indicated that there was no significant difference in the relative effect of BMP-2 on radiographic fusion between different operative approaches or doses. The adverse effects related to the use of BMP-2 included heterotopic bone formation, early bone resorption and graft subsidence.

Conclusions

As compared with autologous bone graft, the use of BMP-2 was associated with a higher proportion of lumbar interbody fusion successes, fewer secondary procedures, and elimination of graft donor site related adverse effects. Limited evidence indicated that there was no significant difference in the effect of BMP-2 on radiographic fusion between different interbody fusion approaches or total dose of BMP-2 applied.

Background

Low back pain causes considerable disability and decreased quality of life, with important health and economic consequences to individuals and societies [1]. Most patients with low back pain are managed by non-surgical interventions including rehabilitation, physiotherapy, psychological and pharmacological therapies [2]. However, in some patients surgical stabilisation of the spine is required. In fact, the rate of surgical interventions, specifically spinal fusion procedures, has been increasing in many countries [3-5]. To achieve a stable union of disc segments in spinal fusion surgery, the conventional technique is to use autologous bone graft (ABG), with related donor-site adverse effects or complications [6]. In addition, the reported rate of non-union by using ABG for spinal fusion ranges from 5% to 45% [7].

Bone morphogenetic proteins (BMPs) are part of the transforming growth factor beta (TGF-beta) super family, which have an important role in bone and cartilage formation, fracture healing and repair of other musculoskeletal tissues [8]. Currently, two BMP products are approved for the use in spinal fusion: BMP-2 (INFUSE, Medtronic) for anterior lumbar interbody fusion in adult patients, and BMP-7 (OP-1 Putty, Stryker) for revision inter-transverse lumbar fusion in compromised patients [9]. Since the approval of BMP-2 in 2002 and BMP-7 in 2003 by the US FDA, the usage of BMP in all spinal fusions in the United States has increased from 0.69% in 2002 to 24.89% in 2006 [10].

A meta-analysis published in 2008 included only trials of BMP in posterolateral fusion of lumbar spine [11]. A systematic review of osteoinductive bone graft substitutes for lumbar fusion found that the rate of radiographic non-union was reduced with the use of BMP-2 but not with BMP-7 [12]. Randomised controlled trials of BMP for spinal fusion have been more inclusively assessed in a report of health technology assessment (HTA) [9]. Since the publication of the HTA report in 2007, new evidence from more recent clinical trials has been published. Therefore, we conducted an updated meta-analysis of controlled studies to evaluate the clinical effectiveness and safety of BMP-2 versus ABG for lumbar spinal fusion.

Methods

Criteria for inclusion of studies

We included randomised (RCT) or non-randomised (non-RCT) controlled studies that compared BMP-2 and ABG in lumbar interbody fusion surgery. Relevant studies included any adult patients with degenerative disc disease, spondylolisthesis, or chronic low back pain. There were no restrictions in relation to BMP-2 dosage or delivery and surgical approaches. Studies were excluded according to the following exclusion criteria: case series or cohort studies without a control group; the follow-up length was shorter than 6 months; studies that evaluated BMP-7; studies included patients with cervical spinal fusion or spinal deformities; and studies that evaluated BMP2 for posterolateral lumbar fusion (PLF).

Identification of relevant studies

The literature searches were limited to human studies, not restricted by language or publication status. The following electronic databases were searched in March 2010 to identify relevant RCTs: Medline (1995 to March 2010); Embase (1980 to March

2010); The Cochrane Central Register of Controlled Trials (*The Cochrane Library*, current issue); Science Citation Index (SCI) Expanded (1995 to March 2010); Current Controlled Trials metaRegister; and ClinicalTrials.gov register. In November 2010, the literature search was expanded to included non-randomised controlled studies. Details of the search strategies used are shown in Additional File 1.

We also searched the bibliographies of studies retrieved, and contacted the industry (Medtronic) to identify further published or unpublished studies.

Two researchers independently screened the titles and abstracts of search results for reports of potentially eligible studies. Any disagreements regarding eligibility were resolved by discussion. Full text reports of potentially eligible studies were examined using the inclusion criteria described above.

Data extraction and assessment

We extracted the following data from included trials: bibliographic details, study and participant characteristics, interventions evaluated, outcomes measures, and results of the trial. Data extraction was carried out by one researcher and checked by another. Multiple reports of the same study were extracted as one study.

The primary outcome measures considered in this meta-analysis were the proportion of successful radiographic fusion, and the improvement in Oswestry Disability Index (ODI) score from baseline. Other outcomes included secondary procedures, operative parameters, any safety outcomes and adverse events.

The assessment of study quality was conducted by one reviewer and checked by a second reviewer. Any disagreement between the two reviewers was resolved by discussion or by the involvement of a third reviewer. The risk of bias in the included RCTs was assessed using the Cochrane Risk of Bias tool [13]. We used a checklist (Additional File 2) to assess the risk of bias in non-randomised studies.

Data analysis methods

We used odds ratio as the outcome statistic for dichotomous data, and the mean difference for continuous data [14]. Standard deviation (SD) for the Oswestry Disability Index (ODI) scores and other continuous data outcomes was not reported in many relevant studies. SDs for the improvement in ODI scores was estimated from presented plots or reported p values in some studies, and data from a few excluded PLF trials were also used because of lack data from the included studies (see Additional File 3 for methods). The pooled estimate of SD of continuous outcomes based on available data was imputed to studies with missing SDs [15].

Statistical heterogeneity across individual studies was tested using the χ^2 test. We also used I^2 statistic to quantify inconsistency across studies [14, 16]. We used random-effects model by default in meta-analysis. Mantel-Haenszel method was used for dichotomous data, and the inverse-variance method for continuous data. Meta-analysis was carried out using the Review Manager (RevMan) software version 5.0 (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2008).

The estimated overall event rates were presented for main binary outcomes. We estimated the overall event rate in the BMP-2 group (*P2*) based on the pooled *OR* and the event rate in the ABG group (*P1*) using the following formula:

$$P2 = \frac{P1 \times OR}{1 - P1 + P1 \times OR} \quad .$$

If possible, the analyses of data from RCTs were carried out according to the intention-to-treat (ITT) principle [17]. For dichotomous outcomes, patients who were lost to follow-up were considered as having an unwanted event. However, for continuous outcomes, the full ITT analysis was impossible and the available case analysis was conducted using data from patients with adequate follow-ups.

Pre-specified subgroup analysis was used to investigate clinically important diversity across studies, by classifying studies according to types of fusion surgery, BMP-2 dose, and diagnosis of lumbar conditions. We also conducted sensitivity analysis according to study design and risk of bias in trials. Funnel plots for the radiographic fusion and the improvement in ODI are presented. The statistical asymmetry of funnel plots was tested using the recommended methods [18, 19].

Results

The study selection process is summarised in Additional File 4. Five unique RCTs [20-24] and seven non-randomised studies were included [25-31]. One RCT study was only available as an abstract [24]. Excluded studies with reasons are shown in Additional File 5.

Characteristics of included studies

The main characteristics of the included five RCTs and seven non-RCT studies are shown in Additional File 6. The participants of the included studies were mainly patients with degenerative disc disease with low grade spondylolisthesis. The mean age of patients ranged from 40 to 58 years old. Where reported, the percentage of patients with previous lumbar surgery was from 25% to 64%.

A single level of lumbar interbody fusion was performed in all five included RCTs and in four of the seven non-RCT studies [26, 29-31]. Anterior lumbar interbody fusion (ALIF) surgery was performed in four studies [20-22, 26, 29], posterior lumbar interbody fusion (PLIF) surgery in three studies [23, 24, 31], and transforaminal lumbar interbody fusion (TLIF) in three studies [25, 28, 30]. One study used either PLIF or TLIF approach [27].

The total dose of BMP-2 administered ranged from 4 to 12 mg per level. A collagen sponge was generally used as the carrier for BMP-2. BMP-2 was used with additional autologous bone graft (ABG) in four non-RCT studies [24, 25, 27, 28]. Except two studies that used only local ABG [24, 27], iliac crest autologous bone graft (ICABG) was used in the control group. A wide range of different devices or instrumentation was used in the included studies (Additional File 6).

Risk of bias assessment

The results of the assessment of risk of bias are summarised in Additional File 7. By design the risk of bias in RCTs is lower than that in the non-RCT studies. So the risk of bias in the two types of studies is separately assessed.

Of the five RCTs, randomisation sequence generation was clearly adequate in three RCTs [20-22], and allocation concealment was explicitly adequate in only two trials [22, 23]. It is not possible to achieve blinding of treatment personnel or patients, although blinded assessors were used for radiographic fusion outcome measurement in all but one trial [24]. One study was prematurely stopped due to the radiographic bone formation extending outside the disc space [23]. Intention-to-treat analysis could be used for the radiographic fusion outcome in four of the five RCTs. Selective

reporting was considered to be unclear in all RCTs, partly due to inadequate data on Oswestry Disability outcome.

Of the seven non-RCT studies, three studies had prospective designs [25, 27, 31], while the remaining four were retrospective in design. The control group was concurrent with the BMP-2 group in four studies [27, 28, 30, 31], and historical in the other three studies [25, 26, 29]. Three non-RCT studies stated that there were no significant differences between groups [29-31]. One study used analysis of covariance method to adjust for differences between groups, including previous back surgery, preoperative non-narcotic medication use, and baseline lower back pain score [26]. Radiographic outcomes were assessed by independent assessors in two studies [26, 29]. Dropouts or withdrawals were appropriately reported and described by three studies [25, 27, 29].

Four of the five RCTs explicitly received funds from industry [20-23] and it is unclear in one [24]. Of the seven non-RCT studies, only one explicitly received industry funds [26], three stated that no external funds were received [25, 27, 29], and the funding source was unclear in the remaining three non-RCT studies [28, 30, 31].

Radiographic fusion success

The pooled odds ratio (OR) of fusion success for all the studies was 2.02 (95% CI: 1.10 to 3.72) (Figure 1). There was moderate but statistically significant heterogeneity between the studies (I^2=56%, p=0.01). Based on the pooled OR (2.02) and the overall fusion rate in the control group (70.1%), the overall fusion rate in the BMP-2 group was 82.6%.

Because of the small number of studies included, possible subgroup analyses were limited (Table 1). The pooled treatment effect by using data from RCTs was statistically significant greater than that by the non-RCT studies (P=0.03). There were no significant differences between different surgical approaches, or between studies that included only patients with single level fusion and studies that included patients with multiple level fusions. The significant difference between different BMP-2 total dose per fusion level (P=0.005) indicated that BMP-2 4.2 mg per level was the most effective for fusion success (Table 1). However, this finding should be interpreted with great caution because the BMP-2 total dose was unclear in three studies [29-31].

Oswestry score improvement

Data on the improvement in the Oswestry Disability Index score (ODI) was available only from RCTs, except one non-RCT study [26] reported ODI back pain score. The difference in the improvement in the Oswestry score was statistically non-significant (mean difference 2.50, 95% CI: -1.15 to 6.14; P=0.18) (Figure 2).

Secondary revision

There was no significant difference in the outcome of secondary procedures between the RCTs and non-RCT studies (Figure 3). The proportion of post-operative secondary procedures was statistically significantly lower in the BMP-2 group as compared with the ABG group (odds ratio 0.55, 95% CI: 0.32 to 0.97; P=0.04). The overall rate of any secondary revision procedures was 17.5% in the control group and 10.4% in the BMP-2 group.

Operative parameters

Because only one or two non-RCT studies reported operative parameters, the analyses were not separated for RCTs and non-RCT studies. The average mean difference in the operative time length between the BMP-2 and the ABG group was -0.62 hour (95% CI: -1.03 to -0.21, P=0.003) (Figure 4). The pooled mean difference in operative blood loss was -35.9 ml (95% CI: -74.0 to 2.14), which was statistically non-significant (P=0.06). The pooled hospital stay length was shorter in the BMP-2 patients (mean difference -1.03, 95% CI: -1.91 to -0.15).

Safety and adverse events

Heterotopic ossification

A study was terminated after observing excess bone formation outside the disc space in the BMP2 arm, although this excess bone formation was not associated with adverse clinical outcomes [23]. Two studies [24, 28] reported no heterotopic bone formation. Joseph et al reported heterotopic bone formation in 5/24 (20.8%) in the BMP-2 group and 1/12 (8.3%) in the control group [27]. In a study of complications after TLIF with BMP-2, ectopic bone formation was observed in 2 of the 86 patients in the BMP-2 group and none of the 33 patients in the autograft group [30]. Burkus et al (2005) reported 18% of patients in the BMP-2 group had transient localised areas of bone remodelling extending into the vertebral body [22]. It is unclear whether heterotopic ossification has not been investigated or was not reported in the remaining studies.

Graft resorption and subsidence

Authors of a non-RCT study observed “early and aggressive resorption” of the femoral ring allograft after adding BMP-2 to FRA, which may cause fracture and collapse of the graft [29]. However, allograft resorption (osteolysis) and related subsidence were not observed in another non-RCT study that used structural femoral allograft with pedicle screw fixation [25]. In a more recent non-RCT study of complications after TLIF, the rate of vertebral osteolysis between 1 to 5 months after surgery was 5.8% (5/86) in the BMP-2 group and 0% (0/33) in the ABG group [30].

According to unpublished data obtained from the US FDA website [32], graft subsidence occurred in 2.4% (7/288) in the BMP-2 group and 1.4% (2/139) in the control group (P=0.72). The percentage of implant displacement or loosening was 1.7% (5/288) in the BMP-2 group and 0.7% (1/139) in the control group (P=0.67).

Graft site related adverse events

All patients in the ICABG control group experienced donor site pain. The reported mean score for graft site pain ranged from 11.3 to 16.0 (out of 20) at discharge, from 2.6 to 6.5 at 6 months, and from 1.8 to 5.5 at 24 months. Donor site pain was reported by 32% to 60% of patients at 24 months.

Other relevant adverse events

Changes in antibody titres were reported by four RCT studies [20-23]. The included RCT studies did not report any fluid collection related adverse events. In one RCT study, retrograde ejaculation in male patients was observed in 11 of the 140 male patients in the BMP-2 group and 1 in 70 male patients in the control group (P=0.065) (data from unpublished file [32]). A non-RCT study reported that the rate of radiculitis in the BMP-2 group after TLIF surgery was higher in the BMP-2 group

than that in the ABG group (14% vs.3%, P=0.08) [30]. It was suspected that the higher rate of radiculitis in the BMP-2 group may be caused by an inflammatory reaction and/or heterotopic bone formation.

Funnel plot assessment of small study effects

Funnel plots (Figure 5) indicated a visual tendency that smaller studies reported greater effect estimates in favour of BMP2. However, the funnel plot asymmetry was not statistically significant.

Discussion

This most up-to-date meta-analysis found that the use of BMP-2 was associated with a higher rate of lumbar fusion than the use of ABG (82.6% vs. 70.1%; p=0.02). Based on limited evidence from subgroup analyses, there was no statistically significant difference in the relative effect of BMP-2 versus ABG on spinal fusion in terms of operative approaches or total dose of BMP-2.

There was no significant difference in the mean Oswestry score improvements between BMP-2 and control groups (P=0.18). There were fewer secondary revisions in BMP-2 patients versus control (10.4% vs. 17.5%; P=0.04). Analysis also found that the use of BMP-2 was associated with shorter operating times (P=0.003) and hospital stay lengths (P=0.02). There was no statistically significant difference in blood loss (p=0.06).

The main adverse effects specifically related to the use of BMP-2 included heterotopic bone formation, early osteolysis (bone resorption) and significant graft subsidence. According to a recent systematic review of complications related to the use of BMP in spine surgery [33], the pooled rate was 44% for resorption/osteolysis, 8% for heterotopic bone formation, and 29% for inflammatory response to collagen carrier. All these reported adverse effects (including possible formation of fluid collections, initial resorption of surrounding trabecular bone, ectopic bone formation and heterotopic ossification) have already been warned in the Summary of Product Characteristics provided by Wyeth Pharmaceuticals [34]. The common adverse events in the ABG group were associated with autogenous bone harvest such as pain, various injuries during the harvest and infection at the site [6].

The effect size of BMP-2 on fusion success by non-RCT studies was smaller than that by the RCT studies (Figure 1). Both RCTs and non-RCT studies included patients mainly with degenerative disc disease and low grade spondylolisthesis. However, non-RCT studies tended to include patients with more varied diagnoses. Patients included in the non-RCTs were much more likely to be treated at multiple fusion levels, whereas single level fusion was conducted in all RCTs. Non-RCT studies in this review were more likely to include either autograft bone, or another bone filler, in conjunction with BMP-2 compared with RCTs.

Caveats and limitations

Of the 12 studies included, some were extremely small. Overall, there was unclear or high risk of bias in the included studies, particularly in non-RCT studies. Most studies were explicitly industry-sponsored. Some adverse effects mentioned in unpublished data were not clearly reported in the corresponding published papers. Although funnel plots were not statistically significantly asymmetric, publication and

reporting bias could not be completely ruled out. For continuous outcomes (for example, changes in ODI scores), the standard deviation was not reported in many studies.

Results of pre-specified subgroup analyses indicated no statistically significant interactions between any subgroups. Since the number of studies included was small, great caution is warranted in the interpretation of findings from the subgroup analyses.

As compared with ABG, the use of BMP-2 increases the initial cost by about €2,266-€2,970 per spinal fusion in Europe [35]. The higher initial cost associated with the use of BMP-2 may be offset to some extent by shorter operative time, shorter hospital stay and fewer secondary procedures [9, 35, 36]. A previous HTA report concluded that, from the societal perspective, the use of BMP-2 for spinal fusion was unlikely to be cost-effective [9]. More recently, Alt and colleagues analysed individual patient data from a clinical study [26] and observed that patients in the BMP-2 group returned to work earlier than those in the ABG group [35]. Consequently, their economic analysis found that the use of BMP-2 in spinal fusion resulted in overall resource savings to society [35]. However, cost-effectiveness of BMP-2 for spinal fusion may need to be investigated by further independent research.

Conclusions

The use of BMP-2 was associated with a higher rate of radiographic fusion and fewer secondary procedures than the use of ABG in lumbar interbody fusion. There was no significant difference in the mean Oswestry score improvement between the BMP-2 and the ABG group. The use of BMP-2 prevents graft site related adverse effects. Limited evidence indicated that there was no significant difference in the effect of BMP-2 on radiographic fusion between different interbody fusion approaches or total dose of BMP-2 applied.

Competing interests

The work was financially supported by Medtronic, the manufacturer of BMP-2 products.

Authors' contributions

All authors contributed to the development of the review protocol. KG and SD conducted literature search. KG extracted data and conducted quality assessment, and FS checked data. KG and FS analysed data and drafted report. All commented on the final manuscript.

References

1. Manchikanti L, Singh V, Datta S, Cohen SP, Hirsch JA: **Comprehensive review of epidemiology, scope, and impact of spinal pain**. *Pain Physician* 2009, **12**(4):E35-70.

2. van Tulder MW, Koes B, Malmivaara A: **Outcome of non-invasive treatment modalities on back pain: an evidence-based review**. *Eur Spine J* 2006, **15 Suppl 1**:S64-81.

3. Chou R, Baisden J, Carragee EJ, Resnick DK, Shaffer WO, Loeser JD: **Surgery for low back pain: a review of the evidence for an American Pain Society Clinical Practice Guideline**. *Spine (Phila Pa 1976)* 2009, **34**(10):1094-1109.

4. Moorin RE, Holman CD: **The impact of the evolution of invasive surgical procedures for low back pain: a population based study of patient outcomes and hospital utilization**. *ANZ J Surg* 2009, **79**(9):610-618.

5. Weinstein JN, Lurie JD, Olson PR, Bronner KK, Fisher ES: **United States' trends and regional variations in lumbar spine surgery: 1992-2003**. *Spine (Phila Pa 1976)* 2006, **31**(23):2707-2714.

6. Sasso RC, LeHuec JC, Shaffrey C: **Iliac crest bone graft donor site pain after anterior lumbar interbody fusion: a prospective patient satisfaction outcome assessment**. *J Spinal Disord Tech* 2005, **18 Suppl**:S77-81.

7. Samartzis D, Khanna N, Shen FH, An HS: **Update on bone morphogenetic proteins and their application in spine surgery**. *J Am Coll Surg* 2005, **200**(2):236-248.

8. Urist MR, Huo YK, Brownell AG, Hohl WM, Buyske J, Lietze A, Tempst P, Hunkapiller M, DeLange RJ: **Purification of bovine bone morphogenetic protein by hydroxyapatite chromatography**. *Proc Natl Acad Sci U S A* 1984, **81**(2):371-375.

9. Garrison KR, Donell S, Ryder J, Shemilt I, Mugford M, Harvey I, Song F: **Clinical effectiveness and cost-effectiveness of bone morphogenetic proteins in the non-healing of fractures and spinal fusion: a systematic review**. *Health Technol Assess* 2007, **11**(30):1-150, iii-iv.

10. Cahill KS, Chi JH, Day A, Claus EB: **Prevalence, complications, and hospital charges associated with use of bone-morphogenetic proteins in spinal fusion procedures**. *JAMA* 2009, **302**(1):58-66.

11. Papakostidis C, Kontakis G, Bhandari M, Giannoudis PV: **Efficacy of autologous iliac crest bone graft and bone morphogenetic proteins for posterolateral fusion of lumbar spine: a meta-analysis of the results**. *Spine (Phila Pa 1976)* 2008, **33**(19):E680-692.

12. Agarwal R, Williams K, Umscheid CA, Welch WC: **Osteoinductive bone graft substitutes for lumbar fusion: a systematic review**. *J Neurosurg Spine* 2009, **11**(6):729-740.

13. Higgins JPT, Altman DG: **Chapter 8: Assessing risk of bias in included studies.** . In *Cochrane Handbook for Systematic Reviews of Interventions Version 502 (updated September 2009)*. Edited by Higgins J, Green S: The Cochrane Collaboration; 2009.

14. Deeks JJ, Higgins JPT, Altman DG: **Chapter 9: Analysing data and undertaking meta-analysis.** In *Cochrane Handbook for Systematic Reviews of Interventions Version 502 (Updated September 2009)*. Edited by Higgins JPT, Green S: The Cochrane Collaboration 2009.

15. Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N: **Imputing missing standard deviations in meta-analyses can provide accurate results**. *J Clin Epidemiol* 2006, **59**(1):7-10.

16. Higgins JP, Thompson SG, Deeks JJ, Altman DG: **Measuring inconsistency in meta-analyses**. *BMJ* 2003, **327**(7414):557-560.

17. Higgins J, Green S: **Cochrane Handbook for Systematic Reviews of Interventions Version 5.0.2 (updated September 2009)**. Cochrane Collaboration; 2009.

18. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L: **Comparison of two methods to detect publication bias in meta-analysis**. *JAMA* 2006, **295**(6):676-680.

19. Egger M, Davey Smith G, Schneider M, Minder C: **Bias in meta-analysis detected by a simple, graphical test**. *BMJ* 1997, **315**(7109):629-634.

20. Boden SD, Zdeblick TA, Sandhu HS, Heim SE: **The use of rhBMP-2 in interbody fusion cages. Definitive evidence of osteoinduction in humans: a preliminary report**. *Spine (Phila Pa 1976)* 2000, **25**(3):376-381.

21. Burkus JK, Gornet MF, Dickman CA, Zdeblick TA: **Anterior lumbar interbody fusion using rhBMP-2 with tapered interbody cages**. *J Spinal Disord Tech* 2002, **15**(5):337-349.

22. Burkus JK, Sandhu HS, Gornet MF, Longley MC: **Use of rhBMP-2 in combination with structural cortical allografts: clinical and radiographic outcomes in anterior lumbar spinal surgery**. *J Bone Joint Surg Am* 2005, **87**(6):1205-1212.

23. Haid RW, Jr., Branch CL, Jr., Alexander JT, Burkus JK: **Posterior lumbar interbody fusion using recombinant human bone morphogenetic protein type 2 with cylindrical interbody cages**. *Spine J* 2004, **4**(5):527-538; discussion 538-529.

24. Shapiro S, Rodgers R, Sloan R, Altstadt T, Miller J: **A randomized trial of instrumented posterior lumbar interbody fusion using machined cortical wedges/local bone with or without rhBMP2 in the treatment of degenerative lumbar spondylolisthesis with stenosis**. *Neurosurgery* 2005, **57**(2):398.

25. Anand N, Hamilton JF, Perri B, Miraliakbar H, Goldstein T: **Cantilever TLIF with structural allograft and RhBMP2 for correction and maintenance of segmental sagittal lordosis: long-term clinical, radiographic, and functional outcome**. *Spine (Phila Pa 1976)* 2006, **31**(20):E748-753.

26. Burkus JK, Heim SE, Gornet MF, Zdeblick TA: **Is INFUSE bone graft superior to autograft bone? An integrated analysis of clinical trials using the LT-CAGE lumbar tapered fusion device**. *J Spinal Disord Tech* 2003, **16**(2):113-122.

27. Joseph V, Rampersaud YR: **Heterotopic bone formation with the use of rhBMP2 in posterior minimal access interbody fusion: a CT analysis**. *Spine (Phila Pa 1976)* 2007, **32**(25):2885-2890.

28. Mummaneni PV, Pan J, Haid RW, Rodts GE: **Contribution of recombinant human bone morphogenetic protein-2 to the rapid creation of interbody fusion when used in transforaminal lumbar interbody fusion: a preliminary report. Invited submission from the Joint Section Meeting on Disorders of the Spine and Peripheral Nerves, March 2004**. *J Neurosurg Spine* 2004, **1**(1):19-23.

29. Pradhan BB, Bae HW, Dawson EG, Patel VV, Delamarter RB: **Graft resorption with the use of bone morphogenetic protein: Lessons from anterior lumbar interbody fusion using femoral ring allografts and recombinant human bone morphogenetic protein-2**. *Spine* 2006, **31**(10):E277-E284.

30. Rihn JA, Patel R, Makda J, Hong J, Anderson DG, Vaccaro A, Hilibrand A, Albert TJ: **Complications associated with single-level transforaminal lumbar interbody fusion**. *The Spine Journal* 2009, **9**:623-629.

31. Xiao R, Li Q, Tang Z: **[Comparative study of lumbar spondylolisthesis treated by three different materials]**. *Zhongguo xiu fu chong jian wai ke za zhi = Zhongguo xiufu chongjian waike zazhi = Chinese journal of reparative and reconstructive surgery* 2007, **21**(5):453-456.

32. Medtronic: **InFUSE Bone Graft/LT-CAGE Lumbar Tapered Fusion Device - Important Medical Information**. In *FDA - Product labelling: http://wwwaccessdatafdagov/cdrh_docs/pdf/P000058cpdf (Accessed on 14/03/2011)* 2002:1-16.

33. Mroz TE, Wang JC, Hashimoto R, Norvell DC: **Complications related to osteobiologics use in spine surgery: a systematic review**. *Spine (Phila Pa 1976)* 2010, **35**(9 Suppl):S86-104.

34. Wyeth_Pharmaceuticals: **Summary of Product Characteristics (InductOs 12mg) last updated on the eMC 27/10/2010**. The electronic Medicines Compendium (eMC) (http://www.medicines.org.uk/emc/medicine/12864#furtherInfo; accessed 24/02/2011); 2010.

35. Alt V, Chhabra A, Franke J, Cuche M, Schnettler R, Le Huec JC: **An economic analysis of using rhBMP-2 for lumbar fusion in Germany, France and UK from a societal perspective**. *Eur Spine J* 2009, **18**(6):800-806.

36. Carreon LY, Glassman SD, Djurasovic M, Campbell MJ, Puno RM, Johnson JR, Dimar JR, 2nd: **RhBMP-2 versus iliac crest bone graft for lumbar spine fusion in patients over 60 years of age: a cost-utility study**. *Spine (Phila Pa 1976)* 2009, **34**(3):238-243.

Table 1 - Radiographic fusion outcome: subgroup analyses

Subgroups	No. of trials (patients)	Odds ratio (95% CI)	Heterogeneity I^2 (P value)	Subgroup difference
All studies	11 (1374)	2.02 (1.10 to 3.72)	56% (0.01)	
RCTs	4 (491)	2.94 (1.07 to 8.04)	67% (0.03)	P=0.03
Non-RCT studies	7 (843)	1.44 (0.65 to 3.22)	32% (0.18)	
ALIF	5 (860)	1.97 (0.88 to 4.50)	75% (0.003)	P=0.90
PLIF	2 (181)	2.57 (0.42 to 15.59)	58% (0.12)	
TLIF	3 (260)	1.56 (0.24 to 10.16)	34% (0.22)	
BMP-2, ≥4.2 mg	3 (147)	12.17 (1.78 to 83.47)	0% (0.93)	P=0.005
BMP-2, 4.2-8.4 mg	3 (746)	1.34 (0.96 to 1.87)	0% (0.58)	
BMP-2, 8.4-12.0 mg	2 (172)	2.74 (0.20 to 37.08)	74% (0.05)	
BMP-2 dose unclear	3 (269)	1.36 (0.23 to 8.01)	60% (0.08)	
Single level operation only	8 (1160)	1.93 (1.01 to 3.68)	63% (0.008)	P=0.51
With multiple levels	3 (174)	3.32 (0.37 to 29.42)	38% (0.20)	

Figure 1 - Fusion success outcome: meta-analysis of using data from RCTs and non-RCT studies

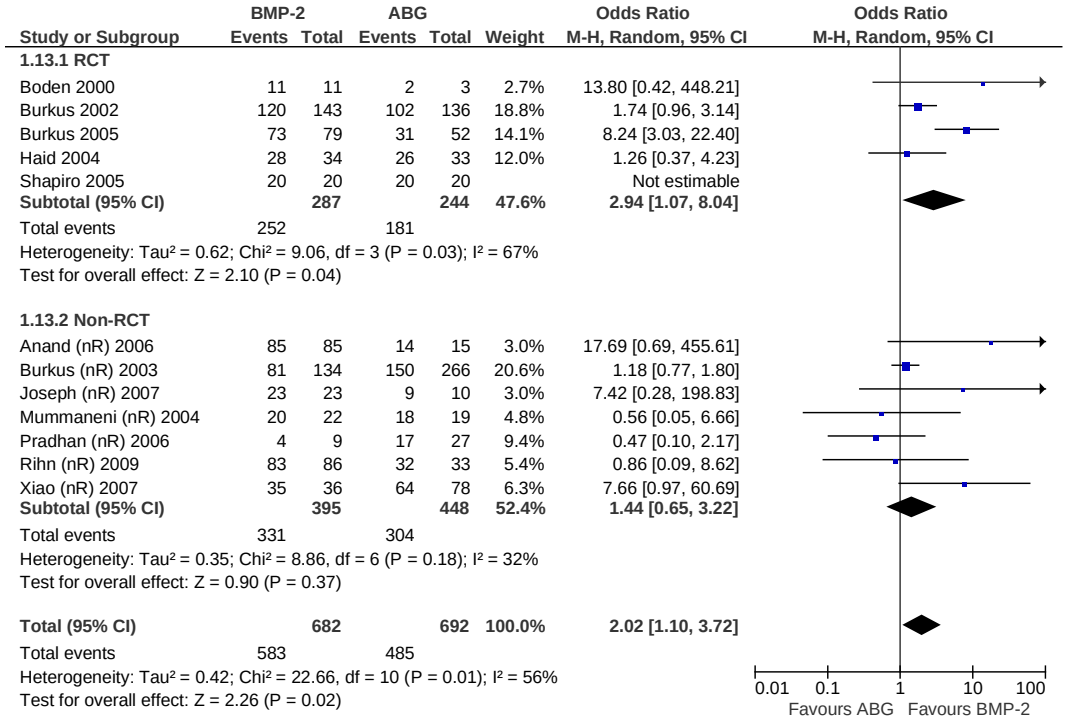


Figure 2 - Improvement in Oswestry Disability Index scores: BMP-2 versus ABG interventions

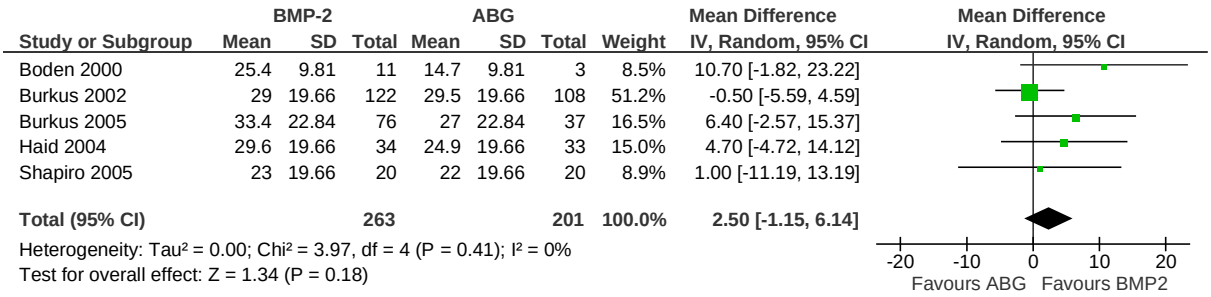


Figure 3 - Proportion of any post-operative secondary procedures: BMP-2 versus ABG interventions

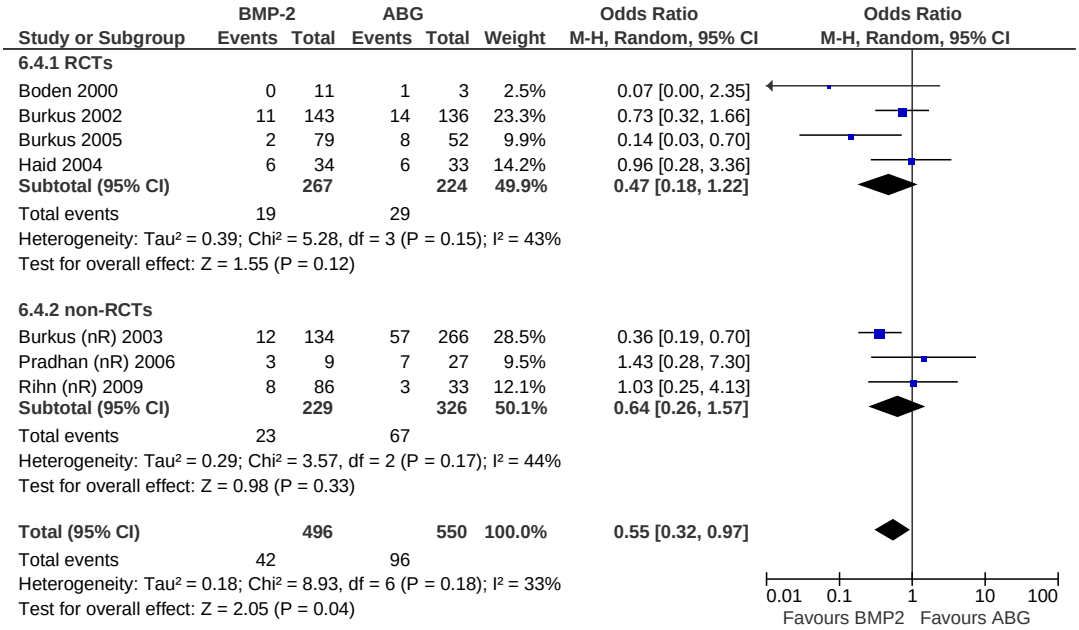
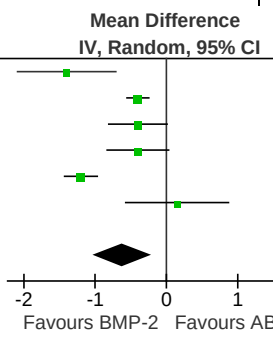


Figure 4 - Results of operative parameters: BMP-2 vs. ABG for lumbar interbody fusion

Mean operative time (hours)

Study or Subgroup	BMP-2 Mean	BMP-2 SD	BMP-2 Total	ABG Mean	ABG SD	ABG Total	Weight	Mean Difference IV, Random, 95% CI
Boden 2000	1.9	0.2	11	3.3	0.6	3	13.2%	-1.40 [-2.09, -0.71]
Burkus 2002	1.6	0.6	143	2	0.7	136	20.2%	-0.40 [-0.55, -0.25]
Burkus 2005	1.4	1.17	79	1.8	1.17	52	17.3%	-0.40 [-0.81, 0.01]
Haid 2004	2.6	0.9	34	3	0.9	33	17.0%	-0.40 [-0.83, 0.03]
Burkus (nR) 2003	1.9	0.9	134	3.1	1.4	265	19.6%	-1.20 [-1.43, -0.97]
Xiao (nR) 2007	2.21	1.78	36	2.06	1.92	78	12.7%	0.15 [-0.57, 0.87]
Total (95% CI)			**437**			**567**	**100.0%**	**-0.62 [-1.03, -0.21]**

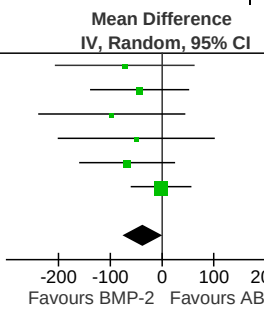
Heterogeneity: Tau² = 0.21; Chi² = 44.34, df = 5 (P < 0.00001); I² = 89%
Test for overall effect: Z = 2.94 (P = 0.003)



Mean operative blood loss (ml)

Study or Subgroup	BMP-2 Mean	BMP-2 SD	BMP-2 Total	ABG Mean	ABG SD	ABG Total	Weight	Mean Difference IV, Random, 95% CI
Boden 2000	95	31	11	167	117	3	8.1%	-72.00 [-205.66, 61.66]
Burkus 2002	109.8	402.2	143	153.1	402.2	136	16.3%	-43.30 [-137.72, 51.12]
Burkus 2005	87.4	402.2	79	184.7	402.2	52	7.3%	-97.30 [-238.07, 43.47]
Haid 2004	322.8	288.3	34	372.7	335.6	33	6.4%	-49.90 [-199.91, 100.11]
Burkus (nR) 2003	146.1	406.2	134	213.6	493	263	17.5%	-67.50 [-158.50, 23.50]
Xiao (nR) 2007	421.67	145.8	36	423.72	142.3	78	44.4%	-2.05 [-59.20, 55.10]
Total (95% CI)			**437**			**565**	**100.0%**	**-35.93 [-74.00, 2.14]**

Heterogeneity: Tau² = 0.00; Chi² = 2.88, df = 5 (P = 0.72); I² = 0%
Test for overall effect: Z = 1.85 (P = 0.06)



Mean post-operative hospital stay length (day)

Study or Subgroup	BMP-2 Mean	BMP-2 SD	BMP-2 Total	ABG Mean	ABG SD	ABG Total	Weight	Mean Difference IV, Random, 95% CI
Boden 2000	2	0.6	11	3.3	1.4	3	16.1%	-1.30 [-2.92, 0.32]
Burkus 2002	3.1	3.8	143	3.3	3.8	136	26.0%	-0.20 [-1.09, 0.69]
Burkus 2005	2.9	3.8	79	3.3	3.8	52	19.6%	-0.40 [-1.73, 0.93]
Haid 2004	3.4	1.3	34	5.2	9.2	33	6.4%	-1.80 [-4.97, 1.37]
Burkus (nR) 2003	1.2	1.1	134	3	3.8	266	31.9%	-1.80 [-2.29, -1.31]
Total (95% CI)			**401**			**490**	**100.0%**	**-1.03 [-1.91, -0.15]**

Heterogeneity: Tau² = 0.57; Chi² = 11.62, df = 4 (P = 0.02); I² = 66%
Test for overall effect: Z = 2.28 (P = 0.02)

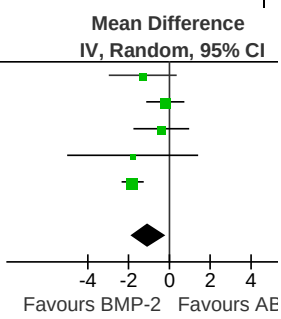
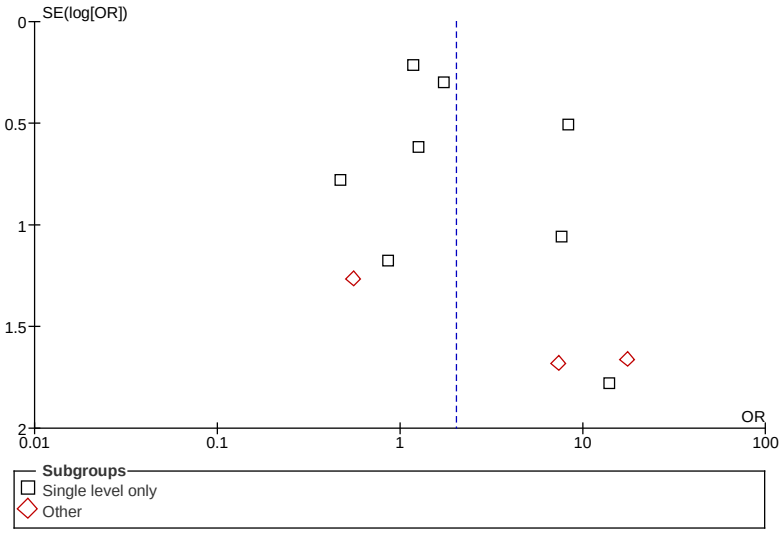
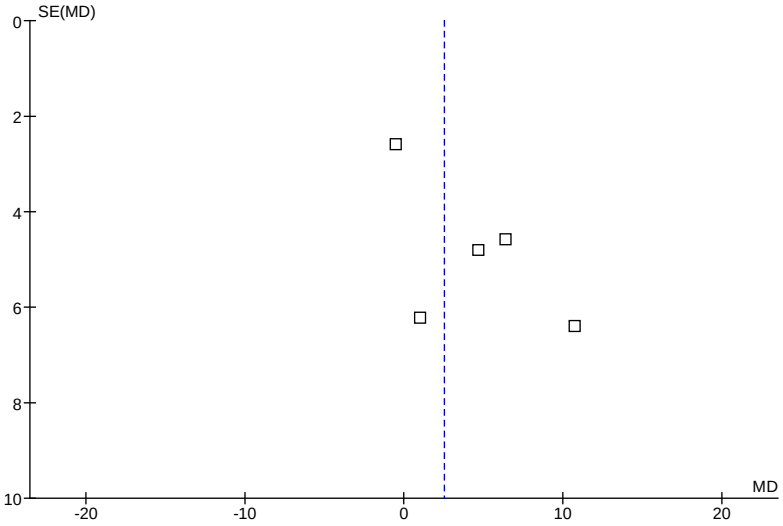


Figure 5 – Funnel plots: fusion success and improvement in ODI scores. The funnel plot asymmetry was not statistically significant (P=0.70 and P=0.12 respectively)

Fusion success outcome



ODI improvement



Additional files

Additional File 1. Literature search strategies

1. Search strategies used to identify RCTs

Medline search strategy (1995 to present)

1. ((bone morphogen$ or osteogen$ or osteoinduct$) adj (protein$ or factor$ or polypeptide$ or polypeptide$)).ti,ab.
2. (BMP2 or BMP-2 or rhBMP2 or rhBMP-2 or rh-BMP2 or rh-BMP-2).ti,ab.
3. (InductOs or InFuse).ti,ab.
4. exp Bone Morphogenetic Proteins/
5. exp Spinal Diseases/
6. exp Spinal Fusion/
7. exp Spinal Injuries/
8. exp Spine/
9. (spine or spinal or thoracic or lumbar or cervical).ti,ab.
10. (non-fusion or nonfusion).ti,ab.
11. fusion.ti,ab.
12. (non-heal$ or nonheal$).ti,ab.
13. (heal or healed or heals or healing).ti,ab.
14. (allograft$ or autograft$ or autogenous or iliac crest).ti,ab.
15. (or/1-4) and (or/5-14)
16. Randomized controlled trial.pt.
17. Controlled clinical trial.pt.
18. Randomized controlled trials/
19. Random allocation/
20. Double blind method/
21. Single blind method/
22. or/16-21
23. Animals/ not Humans/
24. 22 not 23
25. Clinical trial.pt.
26. exp Clinical trials as topic/
27. (clin$ adj25 trial$).ti,ab.
28. ((singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)).ti,ab.
29. Placebos/
30. placebo$.ti,ab.
31. random$.ti,ab.
32. Research design/
33. or/25-32
34. 33 not 23
35. 34 not 24
36. 24 or 35
37. and/15,36

Embase search strategy (1995 to present)

1. ((bone morphogen$ or osteogen$ or osteoinduct$) adj (protein$ or factor$ or polypeptide$ or polypeptide$)).ti,ab.
2. (BMP2 or BMP-2 or rhBMP2 or rhBMP-2 or rh-BMP2 or rh-BMP-2).ti,ab.
3. (InductOs or InFuse).ti,ab.
4. Bone Morphogenetic Protein/ or Bone Morphogenetic Protein 2/
5. exp spine disease/
6. exp spine fusion/
7. exp spine/
8. (spine or spinal or thoracic or lumbar or cervical).ti,ab.
9. (non-fusion or nonfusion).ti,ab.
10. fusion.ti,ab.
11. (non-heal$ or nonheal$).ti,ab.
12. (heal or healed or heals or healing).ti,ab.
13. Healing Impairment/ or Bone Allograft/ or Autograft/
14. (allograft$ or autograft$ or autogenous or iliac crest).ti,ab.

15. (or/1-4) and (or/5-14)
16. exp Randomized Controlled trial/
17. exp Double Blind Procedure/
18. exp Single Blind Procedure/
19. exp Crossover Procedure/
20. Controlled Study/
21. or/16-20
22. ((clinical or controlled or comparative or placebo or prospective$ or randomi#ed) adj3 (trial or study)).ti,ab.
23. (random$ adj7 (allocat$ or allot$ or assign$ or basis$ or divid$ or order$)).ti,ab.
24. ((singl$ or doubl$ or trebl$ or tripl$) adj7 (blind$ or mask$)).ti,ab.
25. (cross?over$ or (cross adj1 over$)).ti,ab.
26. ((allocat$ or allot$ or assign$ or divid$) adj3 (condition$ or experiment$ or intervention$ or treatment$ or therap$ or control$ or group$)).ti,ab.
27. or/22-26
28. or/21,27
29. limit 28 to human
30. and/15,29

Science Citation Index Expanded search strategy (1995 to present)

1. TI=((bone morphogen* osteogen* or osteoinduct*) and (protein* or factor* or polypeptide* or polypeptide))
2. TS=(BMP2 or BMP-2 or rhBMP2 or rhBMP-2 or rh-BMP-2 or rh-BMP2 or InductOs or InFuse)
3. #1 or #2
4. TS=(bone morphogenetic proteins or spinal disease* or spinal fusion or spinal injur* or spine or spinal or thoracic or lumbar or cervical or non-fusion or nonfusion or non-heal* or nonheal* or healing or allograft of autograft or autogenous or iliac crest)
5. #3 and #4
6. TS=(human*)
7. TS=(animal*)
8. #6 not #7
9. #5 and #8

2. Search strategies used to identify non-RCT studies

Medline

1. ((bone morphogen$ or osteogen$ or osteoinduct$) adj (protein$ or factor$ or polypeptide$ or polypeptide$)).ti,ab.
2. (BMP2 or BMP-2 or rhBMP2 or rhBMP-2 or rh-BMP2 or rh-BMP-2).ti,ab.
3. (InductOs or InFuse).ti,ab.
4. exp Bone Morphogenetic Proteins/
5. exp Spinal Diseases/
6. exp Spinal Fusion/
7. exp Spinal Injuries/
8. exp Spine/
9. (spine or spinal or lumbar).ti,ab.
10. (non-fusion or nonfusion).ti,ab.
11. fusion.ti,ab.
12. (non-heal$ or nonheal$).ti,ab.
13. (heal or healed or heals or healing).ti,ab.
14. (allograft$ or autograft$ or autogenous or iliac crest).ti,ab.
15. (or/1-4) and (or/5-14)
16. Clinical trial.pt.
17. exp Clinical trials as topic/
18. (clin$ adj25 trial$).ti,ab.
19. ((singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)).ti,ab.
20. Placebos/
21. placebo$.ti,ab.
22. Research design/
23. or/16-22
24. 15 and 23

25. Humans/
26. 24 and 25
27. limit 26 to yr="1995-Current"

Embase:

1. ((bone morphogen$ or osteogen$ or osteoinduct$) adj (protein$ or factor$ or polypeptide$ or polypeptide$)).ti,ab.
2. (BMP2 or BMP-2 or rhBMP2 or rhBMP-2 or rh-BMP2 or rh-BMP-2).ti,ab.
3. (InductOs or InFuse).ti,ab.
4. Bone Morphogenetic Protein/ or Bone Morphogenetic Protein 2/
5. or/1-4
6. exp spine disease/
7. exp spine fusion/
8. exp spine/
9. (spine or spinal or lumbar).ti,ab.
10. (non-fusion or nonfusion).ti,ab.
11. fusion.ti,ab.
12. (non-heal$ or nonheal$).ti,ab.
13. (heal or healed or heals or healing).ti,ab.
14. Healing Impairment/ or Bone Allograft/ or Autograft/
15. (allograft$ or autograft$ or autogenous or iliac crest).ti,ab.
16. or/6-15
17. 5 and 16
18. ((clinical or controlled or comparative or placebo or prospective$) adj3 (trial or study)).ti,ab.
19. ((singl$ or doubl$ or trebl$ or tripl$) adj7 (blind$ or mask$)).ti,ab.
20. (cross?over$ or (cross adj1 over$)).ti,ab.
21. ((allocat$ or allot$ or assign$ or divid$) adj3 (condition$ or experiment$ or intervention$ or treatment$ or therap$ or control$ or group$)).ti,ab.
22. 18 or 19 or 20 or 21
23. 17 and 22
24. limit 23 to humans and year(1980-current)

Additional File 2 . Checklist for quality assessment of non-RCT studies

The risk of bias in included non-randomised was assessed by one reviewer (KRG) and checked by a second reviewer (FS), using the following checklist [1]:

- Was there a clear description of patient selection, with explicit inclusion/exclusion criteria?
- Was the design prospective or retrospective?
- Was there a concurrent (parallel) control group in the study?
- For controlled studies, were patients comparable in baseline prognostic characteristics?
- For controlled studies, were confounding factors appropriately identified and adjusted?
- Were the study outcomes assessed by independent assessors?
- Were the number of and reasons for drop-outs or withdrawals clearly described?
- Were reports of the study free of suggestions of selective outcome reporting?
- Was the study apparently free of other problems that could put it at a high risk of bias?

Reference:

1. Reeve BC, Deeks JJ, Higgins JP: **Including non-randomized studies**. In *Cochrane Handbook for Systematic Reviews of Interventions Version 502*. Edited by Higgins JP, Green S: The Cochrane Collaboration 2009.

Additional File 3. Methods for imputing missing standard deviations

Furukawas and colleagues demonstrated that imputing missing standard deviations provides accurate results in meta-analyses [1]. Standard deviation (SD) of the mean Oswestry score improvements was only reported in one PLF trial, Glassman 2008 (not included in the paper) [2]. SD was estimated using reported p values in Boden et al 2000 [3] and Burkus et al 2005 [4], or from Figures in Dimar et al 2009 (PLF study not included in the paper) [5] and Dawson et al 2009 (PLF study not included in the paper) [6]. Based on estimated SDs from these 5 trials, we used the following formula to calculate the pooled standard deviations:

$$SD_{poole}d = \sqrt{\frac{\sum[(n_i - 1) \times SD_i^{\,2}]}{\sum(n_i - 1)}}$$

Where *ni* and SD_i *a*re the number of patients and the estimated SD in trial-*i* respectively. Then the pooled SD was imputed to trials that did not report SDs.

References:

1. Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N (2006) Imputing missing standard deviations in meta-analyses can provide accurate results. J Clin Epidemiol 59 (1):7-10

2. Glassman SD, Carreon LY, Djurasovic M, Campbell MJ, Puno RM, Johnson JR, Dimar JR (2008) RhBMP-2 versus iliac crest bone graft for lumbar spine fusion: a randomized, controlled trial in patients over sixty years of age. Spine (Phila Pa 1976) 33 (26):2843-2849

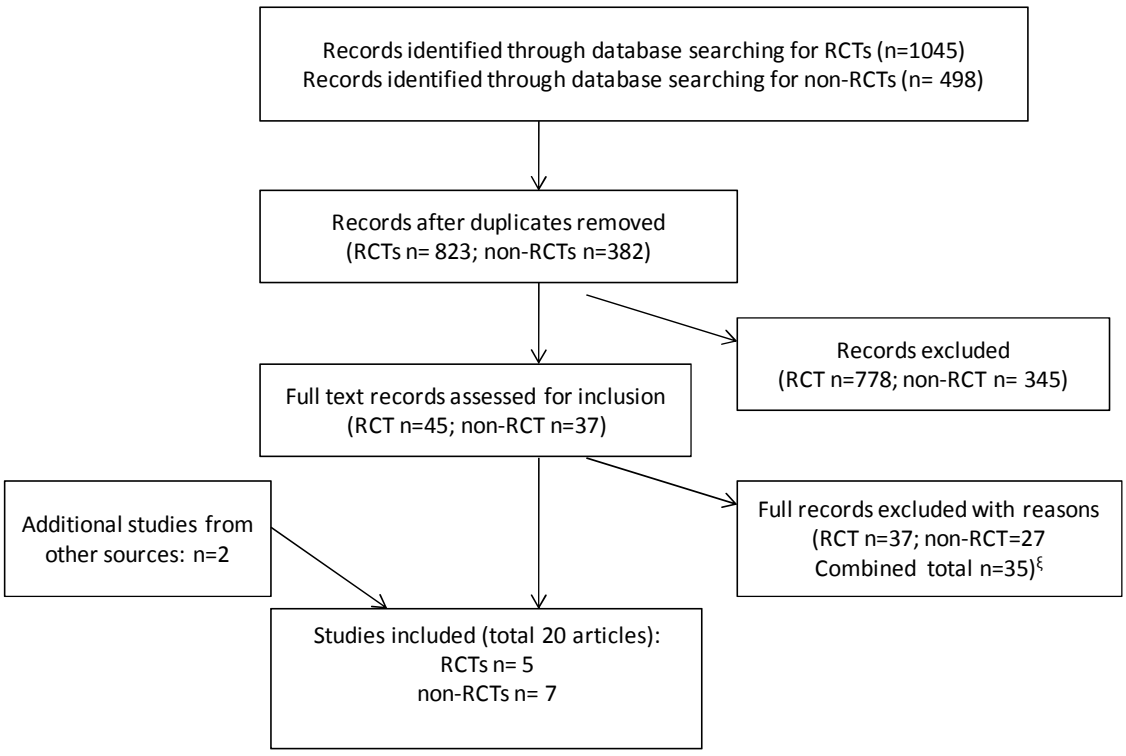
3. Boden SD, Zdeblick TA, Sandhu HS, Heim SE (2000) The use of rhBMP-2 in interbody fusion cages. Definitive evidence of osteoinduction in humans: a preliminary report. Spine (Phila Pa 1976) 25 (3):376-381

4. Burkus JK, Sandhu HS, Gornet MF, Longley MC (2005) Use of rhBMP-2 in combination with structural cortical allografts: clinical and radiographic outcomes in anterior lumbar spinal surgery. J Bone Joint Surg Am 87 (6):1205-1212

5. Dimar JR, 2nd, Glassman SD, Burkus JK, Pryor PW, Hardacker JW, Carreon LY (2009) Clinical and radiographic analysis of an optimized rhBMP-2 formulation as an autograft replacement in posterolateral lumbar spine arthrodesis. J Bone Joint Surg Am 91 (6):1377-1386

6. Dawson E, Bae HW, Burkus JK, Stambough JL, Glassman SD (2009) Recombinant human bone morphogenetic protein-2 on an absorbable collagen sponge with an osteoconductive bulking agent in posterolateral arthrodesis with instrumentation. A prospective randomized trial. J Bone Joint Surg Am 91 (7):1604-1613

Additional File 4. Study selection process



Note: ξ – nine excluded studies for the search of RCTs were included in the search of non-RCTs, and there were many duplicate excluded studies in the two searches.

References of included studies (* the main reference)

RCTs:

Boden 2000 [1]

1.* Boden SD, Zdeblick TA, Sandhu HS, Heim SE (2000) The use of rhBMP-2 in interbody fusion cages. Definitive evidence of osteoinduction in humans: a preliminary report. Spine (Phila Pa 1976) 25 (3):376-381

Burkus 2002 [2-4]

2. * Burkus JK, Gornet MF, Dickman CA, Zdeblick TA (2002) Anterior lumbar interbody fusion using rhBMP-2 with tapered interbody cages. J Spinal Disord Tech 15 (5):337-349

3. Burkus JK, Dorchak JD, Sanders DL (2003) Radiographic assessment of interbody fusion using recombinant human bone morphogenetic protein type 2. Spine (Phila Pa 1976) 28 (4):372-377

4. Gornet M, Burkus J, Dickman C, Zdeblick T (2002) Recombinant human bone morphogenetic protein-2 with tapered cages: a prospective, randomized lumbar fusion study. Proceeding of the NASS 16th Annual Meeting. The Spine Journal (8S-9S)

Burkus 2005 [5-7]

5. * Burkus JK, Sandhu HS, Gornet MF, Longley MC (2005) Use of rhBMP-2 in combination with structural cortical allografts: clinical and radiographic outcomes in anterior lumbar spinal surgery. J Bone Joint Surg Am 87 (6):1205-1212

6. Burkus JK, Sandhu HS, Gornet MF (2006) Influence of rhBMP-2 on the healing patterns associated with allograft interbody constructs in comparison with autograft. Spine (Phila Pa 1976) 31 (7):775-781

7. Burkus JK, Transfeldt EE, Kitchel SH, Watkins RG, Balderston RA (2002) Clinical and radiographic outcomes of anterior lumbar interbody fusion using recombinant human bone morphogenetic protein-2. Spine (Phila Pa 1976) 27 (21):2396-2408

Haid 2004 [8,9]

8. * Haid RW, Jr., Branch CL, Jr., Alexander JT, Burkus JK (2004) Posterior lumbar interbody fusion using recombinant human bone morphogenetic protein type 2 with cylindrical interbody cages. Spine J 4 (5):527-538; discussion 538-529

9. Alexander J, Branch CJ (2002) Recombinant human bone morphogenic protSD stem-2 in a posterior lumbar interbody fusion construct: 2 year-clinical and radiologic outcomes. Spine Journal 2 (5):S48

Shapiro 2005 [10]

10. * Shapiro S, Rodgers R, Sloan R, Altstadt T, Miller J (2005) A randomized trial of instrumented posterior lumbar interbody fusion using machined cortical wedges/local bone with or without rhBMP2 in the treatment of degenerative lumbar spondylolisthesis with stenosis. Neurosurgery 57 (2):398

Non-RCT studies

Anand 2006 [11]

11. * Anand N, Hamilton JF, Perri B, Miraliakbar H, Goldstein T (2006) Cantilever TLIF with structural allograft and RhBMP2 for correction and maintenance of segmental sagittal lordosis: long-term clinical, radiographic, and functional outcome. Spine (Phila Pa 1976) 31 (20):E748-753

Burkus 2003 [12-14]

12. * Burkus JK, Heim SE, Gornet MF, Zdeblick TA (2003) Is INFUSE bone graft superior to autograft bone? An integrated analysis of clinical trials using the LT-CAGE lumbar tapered fusion device. J Spinal Disord Tech 16 (2):113-122

13. Kleeman TJ, Ahn UM, Talbot-Kleeman A (2001) Laparoscopic anterior lumbar interbody fusion with rhBMP-2: A prospective study of clinical and radiographic outcomes. Spine 26 (24):2751-2756

14. Burkus JK, Gornet MF, Schuler TC, Kleeman TJ, Zdeblick TA (2009) Six-year outcomes of anterior lumbar interbody arthrodesis with use of interbody fusion cages and recombinant human bone morphogenetic protein-2. J Bone Joint Surg Am 91 (5):1181-1189

Joseph 2007 [15]

15. * Joseph V, Rampersaud YR (2007) Heterotopic bone formation with the use of rhBMP2 in posterior minimal access interbody fusion: a CT analysis. Spine (Phila Pa 1976) 32 (25):2885-2890

Mummaneni 2004 [16]

16. * Mummaneni PV, Pan J, Haid RW, Rodts GE (2004) Contribution of recombinant human bone morphogenetic protein-2 to the rapid creation of interbody fusion when used in transforaminal lumbar interbody fusion: a preliminary report. Invited submission from the Joint Section Meeting on Disorders of the Spine and Peripheral Nerves, March 2004. J Neurosurg Spine 1 (1):19-23

Pradhan 2006 [17]

17. * Pradhan BB, Bae HW, Dawson EG, Patel VV, Delamarter RB (2006) Graft resorption with the use of bone morphogenetic protein: Lessons from anterior lumbar interbody fusion using femoral ring allografts and recombinant human bone morphogenetic protein-2. Spine 31 (10):E277-E284

Rihn 2009 [18]

18. * Rihn JA, Patel R, Makda J, Hong J, Anderson DG, Vaccaro A, Hilibrand A, Albert TJ (2009) Complications associated with single-level transforaminal lumbar interbody fusion. The Spine Journal 9:623-629

Xiao 2007 [19] [20]

19. * Xiao R, Li Q, Tang Z (2007) [Comparative study of lumbar spondylolisthesis treated by three different materials]. Zhongguo xiu fu chong jian wai ke za zhi = Zhongguo xiufu chongjian waike zazhi = Chinese journal of reparative and reconstructive surgery 21 (5):453-456

20. Xiao RC, Li NN, Tang ZH, Jiang YC, Li Q, Liu Y, Huang WC (2007) [Bone morphogenetic protein versus iliac bone graft substitute with internal fixation in the treatment of osteoporotic intertrochanteric fracture]. Journal of Clinical Rehabilitative Tissue Engineering Research (21):4077-4080

Additional File 5. Excluded full text articles and reasons

Study	Reason for exclusion
Assiri I, du Plessis S, Hurlbert J, Hu R, Salo P, Whittaker T (2004) A prospective randomized clinical study comparing instrumented lumbar fusion rates of Recombinant Human Bone Morphogenic Protein-2 (rhBMP-2) with autogenous iliac crest bone graft in patients with symptomatic degenerative disc disease. Canadian Journal of Surgery 47 (Suppl 4):7-8	PLF (from a previous review)
Boden SD, Grob D, Damien C: Ne-Osteo bone growth factor for posterolateral lumbar spine fusion: results from a nonhuman primate study and a prospective human clinical pilot study. *Spine* 2004, 29(5):504-514.	PLF
Boden SD, Kang J, Sandhu H, Heller JG (2002) Use of recombinant human bone morphogenetic protein-2 to achieve posterolateral lumbar spine fusion in humans: a prospective, randomized clinical pilot trial: 2002 Volvo Award in clinical studies. Spine (Phila Pa 1976) 27 (23):2662-2673	PLF
Burkus, J.K., *Bone morphogenetic proteins in anterior lumbar interbody fusion: old techniques and new technologies.* Journal of Neurosurgery-Spine, 2004. 1(3): p. 254-260.	Review article
Burkus, J.K., et al., *The effectiveness of rhBMP-2 in replacing autograft: An integrated analysis of three human spine studies.* Orthopedics, 2004. 27(7): p. 723-728.	Review article
Dawson E, Bae HW, Burkus JK, Stambough JL, Glassman SD (2009) Recombinant human bone morphogenetic protein-2 on an absorbable collagen sponge with an osteoconductive bulking agent in posterolateral arthrodesis with instrumentation. A prospective randomized trial. J Bone Joint Surg Am 91 (7):1604-1613	PLF
Dimar JR, 2nd, Glassman SD, Burkus JK, Pryor PW, Hardacker JW, Carreon LY (2009) Clinical and radiographic analysis of an optimized rhBMP-2 formulation as an autograft replacement in posterolateral lumbar spine arthrodesis. J Bone Joint Surg Am 91 (6):1377-1386	PLF
Einhorn TA, Einhorn TA: Clinical applications of recombinant human BMPs: early experience and future development. *Journal of Bone & Joint Surgery - American Volume* 2003, 85-A Suppl 3:82-88.	Review
Glassman SD, Carreon LY, Djurasovic M, Campbell MJ, Puno RM, Johnson JR, Dimar JR (2008) RhBMP-2 versus iliac crest bone graft for lumbar spine fusion: a randomized, controlled trial in patients over sixty years of age. Spine (Phila Pa 1976) 33 (26):2843-2849	PLF
Glassman SD, Dimar JR, 3rd, Burkus K, Hardacker JW, Pryor PW, Boden SD, Carreon LY: The efficacy of rhBMP-2 for posterolateral lumbar fusion in smokers. *Spine* 2007, 32(15):1693-1698.	PLF
Glassman, S.D., et al., *Posterolateral lumbar spine fusion with INFUSE bone graft.* Spine Journal, 2007. 7(1): p. 44-49.	PLF, review
Glassman, S.D., et al., *The efficacy of rhBMP-2 for posterolateral lumbar fusion in smokers.* Spine, 2007. 32(15): p. 1693-8.	PLF, sub-analysis
Hamilton DK, Jones-Quaidoo SM, Sansur C, Shaffrey CI, Oskouian R, Jane JA: Outcomes of bone morphogenetic protein-2 in mature adults: posterolateral non-instrument-assisted lumbar decompression and fusion. *Surgical Neurology* 2008, 69(5):457-461.	PLF

Hamilton, D.K., et al., *Outcomes of bone morphogenetic protein-2 in mature adults: posterolateral non-instrument-assisted lumbar decompression and fusion.* Surgical Neurology, 2008. 69(5): p. 457-461.	PLF, not a control study
Katayama, Y., et al., *Clinical and radiographic outcomes of posterolateral lumbar spine fusion in humans using recombinant human bone morphogenetic protein-2: an average five-year follow-up study.* International Orthopaedics, 2009. 33(4): p. 1061-7.	PLF, not a control study
Kuklo, T.R., M.K. Rosner, and D.W. Polly Jr, *Computerized tomography evaluation of a resorbable implant after transforaminal lumbar interbody fusion.* Neurosurgical Focus, 2004. 16(3): p. E10.	Not a control study
Lanman, T.H. and T.J. Hopkins, *Lumbar interbody fusion after treatment with recombinant human bone morphogenetic protein-2 added to poly(L-lactide-co-D,L-lactide) bioresorbable implants.* Neurosurgical Focus, 2004. 16(3): p. E9.	Not a control study
Lee KB, Taghavi CE, Hsu MS, Song KJ, Yoo JH, Keorochana G, Ngo SS, Wang JC: The efficacy of rhBMP-2 versus autograft for posterolateral lumbar spine fusion in elderly patients. *European Spine Journal*, 19 (6):924-930.	PLF
Lewandrowski KU, Nanson C, Calderon R: Vertebral osteolysis after posterior interbody lumbar fusion with recombinant human bone morphogenetic protein 2: A report of five cases. *Spine Journal* 2007, 7 (5):609-614.	<20 patients
Luhmann, S.J., et al., *Use of bone morphogenetic protein-2 for adult spinal deformity.* Spine, 2005. 30(17 Suppl): p. S110-7.	Spinal deformity
Maeda, T., et al., *Long adult spinal deformity fusion to the sacrum using rhBMP-2 versus autogenous iliac crest bone graft.* Spine, 2009. 34(20): p. 2205-2212.	Spinal deformity
McClellan J, Mulconrey D, Forbes R, Fullmer N: Vertebral bone resorption after transforaminal lumbar interbody fusion with bone morphogenetic protein *Journal of Spinal Disorders and Techniques* 2006, 19:483-486.	Not a control study, <6 months follow-up
Meisel, H.J., et al., *Posterior lumbar interbody fusion using rhBMP-2.* European Spine Journal, 2008. 17(12): p. 1735-44.	Not a control study
Mulconrey, D.S., et al., *Bone morphogenetic protein (RhBMP-2) as a substitute for iliac crest bone graft in multilevel adult spinal deformity surgery: minimum two-year evaluation of fusion.* Spine, 2008. 33(20): p. 2153-9.	Spinal deformity
O'Shaughnessy BA, Kuklo TR, Ondra SL: Surgical treatment of vertebral osteomyelitis with recombinant human bone morphogenetic protein-2. *Spine* 2008, 33 (5):E132-E139.	Osteomyelitis
Ray, C.D., *Spinal interbody fusion with Ray threaded titanium cages.* Techniques in Neurosurgery, 1998. 4(3): p. 235-245.	Not a control study
Rogozinski A, Rogozinski C, Cloud G: Accelerating autograft maturation in instrumented posterolateral lumbar spinal fusions without donor site morbidity. *Orthopedics* 2009, 32 (11):809.	PLF
Singh, K., et al., *Use of recombinant human bone morphogenetic protein-2 as an adjunct in posterolateral lumbar spine fusion: a prospective CT-scan analysis at one and two years.[Erratum appears in J Spinal Disord Tech. 2007 Apr;20(2):185 Note: Gill, Sanjitpal [added]]*. Journal of Spinal Disorders & Techniques, 2006. 19(6): p. 416-23.	PLF, not a control study

Smoljanovic T, Bojanic I, Delimar D: Adverse effects of posterior lumbar interbody fusion using rhBMP-2. *European Spine Journal* 2009, 18(6):920-923; author reply 924.	Letter
Smoljanovic, T., S. Vukicevic, and M. Pecina, *Bone morphogenetic protein and fusion.* Journal of Neurosurgery-Spine, 2007. 6(4): p. 378-379.	Not a control study
Solanki G, Vijay S, Chakrapani A, Hendriksz C: Progressive spinal instability in MPS - Use of recombinant human bone morphogenetic protein-2 (RhBMP-2) to augment posterior spine fusion. *Journal of Inherited Metabolic Disease* 2008, 31:435P.	Cervical spine
Stambough JL, Clouse EK, Stambough JB: Instrumented one and two level posterolateral fusions with recombinant human bone morphogenetic protein-2 and allograft: A computed tomography study. *Spine,* 35(1):124-129.	PLF
Vaidya R: Transforaminal interbody fusion and the "off label" use of recombinant human bone morphogenetic protein-2. *Spine Journal* 2009, 9(8):667-669.	Commentary
Vaidya, R., et al., *Complications in the use of rhBMP-2 in PEEK cages for interbody spinal fusions.* Journal of Spinal Disorders & Techniques, 2008. 21(8): p. 557-62.	Not a control study of BMP2 vs. ABG
Villavicencio, A.T., et al., *Safety of transforaminal lumbar interbody fusion and intervertebral recombinant human bone morphogenetic protein-2.* Journal of Neurosurgery Spine, 2005. 3(6): p. 436-43.	Not a control study
Wang JC, Haid RW, Jr., Miller JS, Robinson JC, Wang JC, Haid RW, Jr., Miller JS, Robinson JC: Comparison of CD HORIZON SPIRE spinous process plate stabilization and pedicle screw fixation after anterior lumbar interbody fusion. Invited submission from the Joint Section Meeting On Disorders of the Spine and Peripheral Nerves, March 2005. *Journal of Neurosurgery Spine* 2006, 4(2):132-136.	Not a control study of BMP2 vs. ABG; <6 months follow-up

Additional File 6. Results of assessment of risk of bias in included studies

Risk of bias in randomised controlled trials (RCTs)

Study	Sequence generation	Allocation concealment	Blinding	Incomplete outcome data	Selective reporting	Possible other bias
Boden 2000	Adequate	Unclear	Only for fusion outcome	ITT for fusion outcome	Unclear	Industry supported
Burkus 2002	Adequate	Unclear	Only for fusion outcome	ITT for fusion outcome	Unclear	Industry supported
Burkus 2005	Adequate	Adequate	Only for fusion outcome	ITT for fusion outcome	Unclear	Industry supported
Haid 2004	Unclear	Adequate	Only for fusion outcome	ITT for fusion outcome	Unclear	Industry supported; early trial stop
Shapiro 2005	Unclear	Unclear	Unclear	Unclear	Unclear	Only an abstract, inadequate information

Risk of bias in non-RCT studies

Study	Design (control)	Explicit inclusion /exclusion criteria	Baseline comparability of patients	Confounding factors considered	Independent outcome assessors	Dropouts described	Selective reporting	Funding source
Anand 2006	Prospective (historical)	No	Unclear	No	No	Yes	Unclear	No external funds
Burkus 2003	Retrospective (historical)	Yes	Some differences between groups	Yes	For fusion outcome	No, unknown reasons	Unclear	Industry
Joseph 2007	Prospective (parallel)	No	Unclear	No	Unclear	Yes	Unclear	No external funds
Mummaneni 2004	Retrospective (parallel)	No	Unclear	No	No	No, unknown reasons	Unclear	Unclear
Pradhan 2006	Retrospective (historical)	Yes	Stated no significant differences	No	For fusion outcome	Yes	Unclear	No external funds
Rihn 2009	Retrospective (parallel)	Yes	Stated no significant differences	No	No	Not separated by group	Unclear	Unclear
Xiao 2007	Prospective (parallel)	No	Stated no significant differences	No	No	No	Unclear	Unclear

Additional file 7. Main characteristics of the included studies

Study	Design, sample size: total (BMP-2 , control)	Length of follow-up	Surgical procedure, clinical indications, and Previous spinal surgery	Mean age (range): BMP-2 vs. Control	Male%: BMP-2 vs. Control	BMP-2 group (dose and carrier)	Control group	Implant, and other Instrumentation
Randomised controlled trials								
Boden 2000 {Boden, 2000 #54}	RCT 14 (11, 3)	24 months	ALIF (open or laparoscopic), all single level Single level degenerative disc disease or grade 1 spondylolisthesis Previous surgery: NR	42.5 (30-62) vs. 40.2 (38-42)	45.5% vs. 66.7%	BMP-2 (total: 4mg) ACS	ICABG	NOVUS LT tapered interbody fusion device
Burkus 2002 {Burkus, 2002 #53}	RCT 279 (143, 136)	24 months	ALIF (open), all single level Single level degenerative disc disease Previous surgery: NR	43.3 vs. 42.3	54.5% vs. 50.0%	BMP-2 (total: 4.2 to 8.4mg) ACS (INFUSE)	ICABG	LT-CAGE lumbar Tapered Fusion Device
Burkus 2005 {Burkus, 2005 #36}	RCT 131 (79, 52)	24 months	ALIF, all single level Degenerative disc disease or grad 1 spondylolisthesis Previous surgery: 37% vs. 33%	40.2 vs. 43.6	40.5% vs. 36.5%	BMP-2 (total: 8.4 to 12mg) ACS (INFUSE)	ICABG	Both groups received a pair of threaded cortical allograft dowels
Haid 2004 {Haid, 2004 #41}	RCT 67 (34, 33)	24 months	PLIF, all single level Single level degenerative disc disease Previous surgery: 35% vs. 39%	46.3 (25.8-66.1) vs. 46.1 (28.5-70.9)	50% vs. 45.5%	BMP-2 (total: 4.0 to 8.0mg) ACS (INFUSE)	ICABG	INTER FIX interbody cages
Shapiro 2005 {Shapiro, 2005 #116}	RCT 40 (20, 20)	12 months	PLIF, all single level Severe spinal stenosis & grade 1 spondylolisthesis Previous surgery: NR	NR	NR	BMP-2 (total: 4.2mg) ACS +local ABG	Local ABG	Machined cortical wedges. Used screw/rod

Study	Design, sample size: total (BMP-2 , control)	Length of follow-up	Surgical procedure, clinical indications, and Previous spinal surgery	Mean age (range): BMP-2 vs. Control	Male%: BMP-2 vs. Control	BMP-2 group (dose and carrier)	Control group	Implant, and other Instrumentation
Non-randomised controlled studies								
Anand 2006 {Anand, 2006 #27}	Non-RCT (prospective) 100 (85, 15)	Mean 30 (24-47) months	TLIF (laparoscopic), 1.24 level per patient. Degenerative disc disease with or without spondylolisthesis. Previous surgery: all 64%.	All patients: 52 (range 29-76)	All patients 42%	BMP-2 (total 4.2mg) ACS + local ABG	ICABG + local ABG	Structural femoral allograft ring, and pedicle screw fixation
Burkus 2003 {Burkus, 2003 #48}	Non-RCT (Retrospective) 400 (134, 266)	24 months	ALIF (laparoscopic), all single level. Single level degenerative disc disease or low-grade (grade I) spondylolisthesis. Previous surgery: 24.6% vs. 41.4%	42.4±10.5 vs. 40±9.6	42.5% vs. 53.4%	BMP-2 (4.2-8.0mg). ACS	ICABG	NOVUS LT Cage
Joseph 2007 {Joseph, 2007 #68}	Non-RCT (Retrospective) 35 (23, 12)	Mean 26 (23-30) months	PLIF or TLIF (laparoscopic), 1.09 levels per patient. Spondylolisthesis or degenerative disc. Previous surgery: NR	All patients: 47.9 (range 22-69)	All patients: 61%	BMP-2 (total 4.2mg) ACS +local ABG	Local ABG (morcellised)	Non-threaded or Boomerang cage and transpendicular pedicle screw fixation
Mummaneni {Mummaneni, 2004 #42}	Non-RCT (Retrospective) 43 (24, 19)	9 (3-19) months	TLIF (open or laparoscopic), 1.05 levels per patient. Degenerative disc disease, Grade 1-2 spondylolisthesis. Previous surgery: 40% vs. 32%	58 (range 33-76) vs. 49 (range 33-64)	68% vs. 47%	BMP-2 (total 8.4mg) ACS + ICABG or local ABG	ICABG	Titanium paramesh or polyetherketone boomerang shaped cage. All with pedicle screw fixation

Study	Design, sample size: total (BMP-2 , control)	Length of follow-up	Surgical procedure, clinical indications, and Previous spinal surgery	Mean age (range): BMP-2 vs. Control	Male%: BMP-2 vs. Control	BMP-2 group (dose and carrier)	Control group	Implant, and other Instrumentation
Pradhan 2006 {Pradhan, 2006 #135}	Non-RCT (Retrospective) 36 (9, 27)	24 (23-55) months	ALIF, all single level Single level degenerative disc disease. Patients with previous surgery excluded: 0%	51.2 vs. 53.4 (range NR)	30.0% vs. 18.5%	BMP-2 (dose NR) ACS	ICABG	Stand alone femoral ring allograft
Rihn 2009 {Rihn, 2009 #137}	Non-RCT (Retrospective) 119 (86, 33)	19 (10-35) months	TLIF, all single level Degenerative disc disease, spondylolisthesis . Previous surgery: 37%	All patients: 47.4 (range NR)	All patients: 52.9%	BMP-2 (dose NR) Some +Duraseal (n=37)	ICABG	Interbody cage
Xiao 2007 {Xiao, 2007 #146}	Non-RCT (Prospective) 114 (36, 78)	15 (13-30) months	PLIF, all single level Spondylolisthesis (Grade 1-3) Previous surgery: NR	43 (range 33-58) vs. A: 43 (range 32-61); B: 42 (range 35-56)	37.8% vs. 42.9% or 36.1%	BMP-2 (dose NR) Cage +Artificial bone material	A: ICABG only (n=42) B: Cage +ICABG (n=36)	Titanium fusion cage, artificial bone material, and pedicle screw fixation

Notes:

- ABG – Autologous bone graft
- ACS – Absorbable collagen sponge
- ALIF – anterior lumbar interbody fusion.
- ICABG – Iliac crest autologous bone graft
- NR – not reported.
- PLIF – posterior lumbar interbody fusion.
- TLIF – Transforaminal lumbar interbody fusion.