



# A Meta Analysis of Lumbar Spinal Fusion Surgery Using Bone Morphogenetic Proteins and Autologous Iliac Crest Bone Graft

Haifei Zhang<sup>1</sup>, Feng Wang<sup>1</sup>, Lin Ding<sup>2</sup>, Zhiyu Zhang<sup>1</sup>, Deri Sun<sup>1</sup>, Xinmin Feng<sup>1</sup>, Jiuli An<sup>1</sup>, Yue Zhu<sup>3\*</sup>

<sup>1</sup> Department of Orthopedics, The fourth Affiliated Hospital, China Medical University, Shenyang, China, <sup>2</sup> China Medical University Computer Center, Shenyang, China, <sup>3</sup> Department of Orthopedics, The first Affiliated Hospital, China Medical University, Shenyang, China

## Abstract

**Background:** Bone morphogenetic protein (BMPs) as a substitute for iliac crest bone graft (ICBG) has been increasingly widely used in lumbar fusion. The purpose of this study is to systematically compare the effectiveness and safety of fusion with BMPs for the treatment of lumbar disease.

**Methods:** Cochrane review methods were used to analyze all relevant randomized controlled trials (RCTs) published up to nov 2013.

**Results:** 19 RCTs (1,852 patients) met the inclusion criteria. 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,  $p=0.0002$ ). Significant difference was found in the operating time (MD–0.32; 95% CI–0.55, –0.08;  $P=0.009$ ), but no significant difference was found in the blood loss, the hospital stay, patient satisfaction, and work status.

**Conclusion:** 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.

**Citation:** Zhang H, Wang F, Ding L, Zhang Z, Sun D, et al. (2014) A Meta Analysis of Lumbar Spinal Fusion Surgery Using Bone Morphogenetic Proteins and Autologous Iliac Crest Bone Graft. PLoS ONE 9(6): e97049. doi:10.1371/journal.pone.0097049

**Editor:** Mohammed Elsalanty, Georgia Regents University, College of Dental Medicine, United States of America

**Received:** August 26, 2013; **Accepted:** April 8, 2014; **Published:** June 2, 2014

**Copyright:** © 2014 Zhang et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The authors have no support or funding to report.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: zhuyuedr@163.com

## Introduction

Autogenous iliac crest bone graft (ICBG) is considered the gold standard graft material for lumbar fusion, but there are several serious shortcomings in performing lumbar arthrodesis with ICBG, including donor-site morbidity and relatively high frequency of nonunion. Additionally, the amount and quantity of autogenous bone graft are limited and may be insufficient, particularly in arthrodesis over multiple segments [1]. In an effort to decrease the reliance on autograft, bone morphogenetic protein (BMPs) which Urist first described in 1965 had been utilized to supplement or replace the bone graft in spinal fusion surgery, but mass production of this molecule became feasible after the sequencing of multiple BMP genes in the 1990s [2,3]. Human BMP is now produced on a large scale using recombinant techniques. Since the FDA, investigational device exemption for rhBMP-2 in 1996 and for rhBMP-7 in 2001, both BMPs have been under clinical investigation in various trials. So, we conducted this meta-analysis to assess the effectiveness and safety of BMPs compared with ICBG in lumbar fusion.

## Materials and Methods

### Literature Search

A protocol was developed in advance of conducting this meta-analysis following the Cochrane Back Review Group guidelines [6]. Updating to November 2013, the relevant RCTs in all languages were identified through computer and other research methods. The sources of computer searching include PubMed, The Cochrane Central Register of Controlled Trials (CENTRAL), Ovid MEDLINE and EMBASE, CINAHL, the China Biological Medicine Database (CBM), International Clinical Trials Registry Platform (ICTRP), Current Controlled Trials, ClinicalTrials.gov. Other searching methods include screening references listed in relevant systematic reviews and identified RCTs, and searching abstracts of relevant meetings, and personal communication with content experts in the field and with authors of identified RCTs. Key words that have been used for researching are lumbar degenerative disease (LDD), low back pain, lumbar fusion, bone morphogenetic protein-2, bone morphogenetic protein-7, osteogenic protein-1, and randomized controlled trial.

## Study Eligibility Criteria

All RCTs comparing the BMPs to ICBG for the treatment of LDD were identified in this study. Patients older than eighteen years of age with systematic LDD were included in the review. Articles were regarded eligible if they met the following inclusion criteria: the target population comprised adult patients suffering from degenerative conditions of the lumbar spine requiring fusion; the main intervention was lumbar fusion using BMPs as a substitute to ICBG; each potentially eligible study included a comparison group of patients in whom ICBG was used as the only biologic enhancement of the fusion process. Articles were excluded if they reported on patient populations with any of the following characteristics: spinal deformities in adolescents, fractures of the spinal column, spondylolisthesis classified as higher than Meyerding Grade 2, a regular postoperative regimen of pharmaceutical agents that potentially could interfere with fusion (such as steroids or chemotherapy agents).

The trial selection process was based on a first phase of title and abstract screening followed by a second phase of eligibility evaluation from the full-text format. Both actions were performed by two reviewers and checked by the principal reviewer. The observed percentage agreement between the reviewers for the assessment of inclusion was calculated using the  $\kappa$  test [4,5]. Disagreements were resolved by discussion.

## Risk of Bias Assessment and Evaluation of Validity

The risk of bias (RoB) and methodological quality was assessed in duplicate using the 12 criteria recommended by the Cochrane Back Review Group and evaluated independently by two review investigators [6,7]. A study with a low RoB was defined as one fulfilling six or more of the criteria items, which is supported by empirical evidence, and with no fatal flaw, which is defined as those studies with (1) a dropout rate greater than 50% at the first and subsequent follow-up measurements or (2) statistically and clinically relevant important baseline differences for one or more primary outcomes indicating unsuccessful randomization. The quality of the evidence related to the estimation of lumbar fusion with BMPs and ICBG followed the suggestions of the GRADE Working Group by adopting the use of GradePro (version 3.6).

## Data Extraction

The data were extracted from included reports independently by two reviewers, and further discussions were done to deal with the disagreements. The data extracted included the following categories: the participant characteristics, the number of participants, and the loss to follow-up; study characteristics; intervention details; the primary and the secondary outcomes. The primary outcomes included: (1) the solid fusion rate, (2) clinical outcomes, (3) complications, and (4) the reoperation rate. The secondary outcomes included: (1) the operation time and blood loss, and hospital stay, (2) patient satisfaction with the treatment, (3) work status and return to work rate.

## Assessment of Heterogeneity

Heterogeneity was explored in two manners, informally by vision (eye-ball test) and formally tested by the  $Q$ -test (chi-square) and  $I^2$ ; however, the decision regarding heterogeneity was dependent on  $I^2$ . Substantial heterogeneity is defined as  $\geq 50\%$ , and where necessary, the effect of the interventions is described if the results are too heterogeneous.

## Assessment of Clinical Relevance

Two reviewers independently assessed the clinical relevance of included studies according to 5 questions that were recommended by the Cochrane Back Review Group [6]. Each question was scored positive (+) if the clinical relevance item was met, negative (−) if the item was not met, and unclear (?) if data were not available to answer the question. A 20% improvement in pain scores and a 10% improvement in functioning outcomes were considered to be clinically important.

## Measures of Treatment Effect

Attempts were made to statistically pool the data of homogeneous studies in order to obtain the primary and the secondary outcomes. The results were expressed in terms of risk ratio (RR) and a 95% confidence interval (95% CI) for dichotomous outcomes, and in terms of mean difference (MD) and 95% CI for continuous outcomes. When the same continuous outcomes are measured in different scales, standardized mean difference (SMD) and 95% CI are calculated. If in some studies outcomes are shown as dichotomous data while in the other studies expressed as continuous data, RRs would be expressed as SMD to allow dichotomous and continuous data to be pooled together [6]. Collected data were checked and entered into the computer by the two reviewers. A random-effects model was used in this meta-analysis [6,8]. We performed a sensitivity analysis for the measured effects omitting studies with low methodological quality which may largely influence the clinical results. Funnel plot and statistic tests (Egger's test and Begg's test) were used to explore potential publication bias [9–11]. To assess the stability in the overall result if publication bias existed, we corrected the summary results by the trim and fill method [12,13]. RevMan software (version 5.1.0) and the R project (version 3.0.1) were used for data analysis.

## Results

### Search Results

The primary search identified 244 records, and 166 publications were immediately excluded based on titles and abstracts. From the potentially relevant 78 publications, 59 were omitted according to the inclusion criteria. Finally, 19 trials [14–32] were included in the meta-analysis (Fig. 1). The  $\kappa$  statistic for interrater agreement in terms of study eligibility was 0.81.

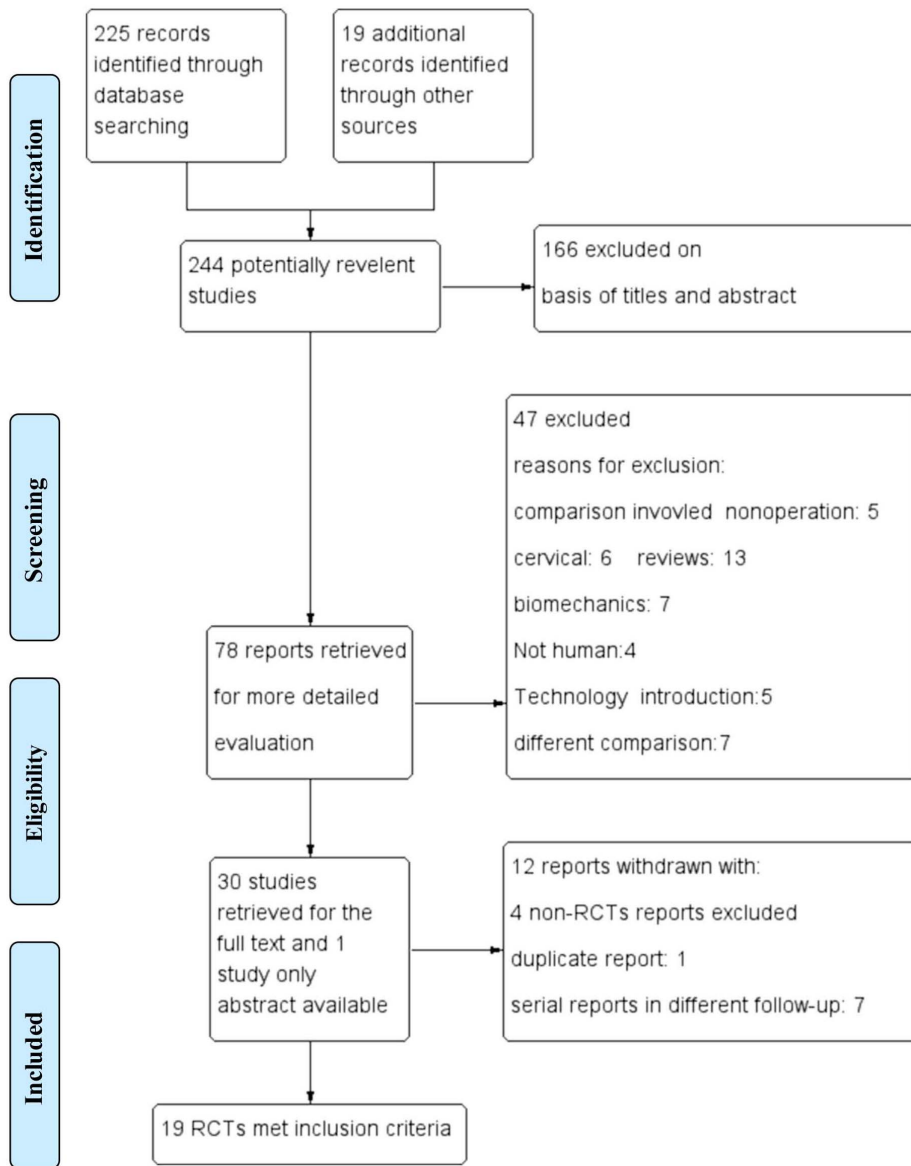
19 RCTs (15 bmp-2, 4 bmp-7) involving 1852 patients were deemed eligible for inclusion, with individual sample sizes ranging from 14 to 463 patients. All the included studies have definite inclusion/exclusion criteria. Those studies recruited patients with a variety of spinal disorders, and surgical treatment involved ALIF fusion, PLIF fusion or PLF fusion (instrumented OR uninstrumented). In one study [15], there were two treatment groups (rhBMP-2/TSRH group and rhBMP-2 only group) and one control group (ICBG/TSRH group). To avoid heterogeneity, rhBMP-2 only group was omitted from this Meta analysis. Characteristics of included studies were presented in Table 1.

### Methodological Quality of the Included Studies

The results of the RoB for the individual studies are summarized in Figure 2. In total, 12 of the 19 trials met the criteria for a low RoB [14–18,23,27–32]. 6 studies have adequate methods of randomization [14,19,23,28,29,31], and only two studies use both an adequate sequence generation and allocation procedure [29,31]. In 9 studies, both randomization and allocation were unclear [15,17,20,22,24,25–27,30].



## PRISMA Flow Diagram



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit [www.prisma-statement.org](http://www.prisma-statement.org).

**Figure 1. PRISMA flow diagram.**  
doi:10.1371/journal.pone.0097049.g001

No reliable studies attempted to blind patients or surgeon because this was impossible due to the nature of the surgery. The lack of blinding was compensated by using blinded observers to assess the fusion outcome in 10 studies [14,16,21,23,27–32]. To prevent any potential bias in surgical technique between the treatment groups, 3 studies [18,31,32] revealed the randomization

at the end of the surgery, just before the graft was needed. So, we considered those studies as blinded care provider.

Most of the studies provided an adequate overview of withdrawals or dropouts and were able to keep these to a minimum for the subsequent follow-up measurements, although only Vaccaro and Burkus conducted long-term follow-up [24,28].

**Table 1.** Overview of included trials.

1 <sup>st</sup> Author, Publication year	Country	Conducted Preoperative diagnosis	Comparisons	Sample size(T/C)	Female (%)	Mean age(year)T/C	Follow-up (month)	Follow-up rate(%)
Boden 2000 [14]	United States	Single-level lumbar DDDspondylosis ≤ grade 1	rhBMP-2/ACS vs. ICBG (ALIF)	11/3	50	42.5/40.2	24	100
Boden 2002 [15]	United States	Single-level lumbar DDDspondylosis ≤ grade 1	rhBMP-2/BCP vs. ICBG (PLF with/without TSRH)	11/5	68.8	57.6/52.9	17	100
Burkus Gornet 2002[16]	United States	Single-level lumbar DDD	rhBMP-2/ACS vs. ICBG.(ALIF)	143/136	47.7	43.3/42.3	24	91.7
Burkus Transfeldt 2002 [17]	United States	Single-level lumbar DDD	rhBMP-2(InFUSE) vs. ICBG.(ALIF)	24/22	60.9	41.5/45.6	24	95.7
Johnsson 2002 [18]	Sweden	L5 spondylosis and vertebral slip ≤50%,	BMP-7 vs. ICBG.(uninstrumented PLF)	10/10	60	42.9/40.4	12	100
Burkus 2003 [19]	United States	degenerative lumbar spondylosis	rhBMP-2/ACS vs. ICBG.(ALIF)	22/20	47.6	41.7/44.2	24	100
Assiri 2004 [20]	Canada	DDD	rhBMP-2 vs. ICBG. (PLF)	8/7	NR	NR	24	NR
Haid 2004 [21]	United States	Single-level lumbar DDDspondylosis ≤ grade 1	rhBMP-2/ACS vs. ICBG.(PLIF)	34/33	52.2	46.3/46.1	24	94
Glassman 2005 [22]	United States	Single-level lumbar DDDspondylosis ≤ grade 1	rhBMP-2/CRM vs. ICBG.(PLF)	38/36	59.5	53/53	12	97
Vaccaro 2005 [23]	United States	single-level degenerative spondylosis and stenosis	BMP-7 vs. ICBG.(uninstrumented PLF)	24/12	55.6	63/66	24	88.9
Burkus 2005 [24]	United States	Single-level lumbar DDD	rhBMP-2/ACS vs. ICBG.(ALIF)	79/52	61.1	40.2/43.6	24	99.2
Dimar 2006 [25]	United States	Single-level lumbar DDDspondylosis ≤ grade 1	rhBMP-2/CRM vs. ICBG.(PLF)	53/45	57.1	50.9/52.7	24	100
Kanayama 2006 [26]	Japan	degenerative spondylosis with stenosis	BMP-7 vs. ICBG.(PLF)	9/10	42.1	70.3/58.7	13–16	100
Glassman 2008 [27]	United States	Single/multilevel lumbar DDD; spondylosis; stenosis; adjacent level fusion	rhBMP-2/ACS vs. ICBG.(PLF)	50/52	68.6	69.9/69.2	24	94.3
Vaccaro 2008 [28]	United States	single-level degenerative spondylosis and stenosis	BMP-7 vs. ICBG.(uninstrumented PLF)	208/87	66.9	68/69	>48	80
Dimar 2009 [29]	United States	Single-level lumbar DDDspondylosis ≤ grade 1	rhBMP-2 matrix vs. ICBG.(PLIF)	239/224	56.2	53.2/52.3	24	89
Dawson 2009 [30]	United States	Single-level lumbar DDDspondylosis ≤ grade 1	rhBMP-2/ACS vs. ICBG.(PLIF)	25/21	58.7	55.9/56.9	24	87
Delawi 2010 [31]	Netherlands	Spondylosis ≤ grade 2	BMP-7 vs. ICBG.(PLF)	18/16	52.9	53/55	12	94.4
Michielssen 2013 [32]	Belgium	Single-level lumbar DDD	rhBMP-2/ACS vs. ICBG.(PLIF)	19/19	57.8	43.2/42.2	12	100

ACS: absorbable collagen sponge carrier; ICBG: autogenous iliac crest bone graft;

BCP: biphasic calcium phosphate; CRM: compression resistant matrix; DDD: degenerative disc disease; NR: no report.

doi:10.1371/journal.pone.0097049.t001

risk of bias in seven trial designs and implementation. As other primary outcomes, overall success and reoperation rate were also rated as low quality because of imprecision and/or reporting bias, and complications were rated as moderate quality on account of imprecision. However, the secondary outcomes were rated as moderate quality, low quality, or very low quality, results from assessing pooled events of patient satisfaction, surgical conditions and work status respectively.

Clinical Relevance

Clinical relevance of included studies was presented in Table 3. The  $\kappa$  statistic for interrater agreement in terms of study eligibility was 0.83. Consensus was reached on all scorings after discussion. The reviewers considered the likely treatment benefits to be worth the potential harms in 13 studies [14,16,17,21–24,25,27–31], and the size of the effect was considered to be clinically important in eight studies [14,15,17,22–24,30,32], and all clinically relevant outcomes were considered to be measured and reported in nine studies [14,15,17,23,25,27,28,31,32]. Most of the included trials described the interventions and treatment settings well enough to enable clinicians to replicate the treatment in clinical practice.

Quantitative Data Synthesis

17 studies [14,15,17–27,29–32] assessed the fusion rate between BMPs and ICBG (610 participants with BMPs and 523 with ICBG), significant differences were found in comparisons (RR: 1.13; 95% CI 1.05–1.23;  $P=0.003$ ). Heterogeneity was obvious during follow-up 24 months,  $I^2=52\%$ . We also have a subgroup analysis. Similar results were obtained by pooled only BMP-2 studies (RR: 1.16; 95% CI 1.06–1.27;  $P=0.001$ ), by contrast, pooled BMP-7 studies have different results (RR: 0.90; 95% CI 0.69–1.17;  $P=0.43$ ). Heterogeneity were moderate or absent in bmp-2 subgroup and bmp-7 subgroup, respectively (BMP-2:  $I^2=62\%$ ; BMP-7:  $I^2=0$ ; Fig. 3). Data for overall success of clinical outcomes were available in 8 studies (431 participants with BMPs and 265 with ICBG) [14–17,21,23,28,30]. No significant difference was found between two groups (RR: 1.04; 95% CI 0.95–1.13;  $P=0.38$ ). There was no significant heterogeneity between trials ( $I^2=2\%$ ; Fig. 4). With regard to complications, we pooled data of 9 trials [15,21,23,27–31] about the frequency of adverse reactions (605 participants with BMPs and 444 with ICBG). The frequency of adverse events or complications was similar in both groups (RR = 0.96; 95% CI 0.85–1.09;  $p=0.54$ ). There was no heterogeneity between the studies ( $I^2=0\%$ ; Fig. 5). The reoperation rate of the BMPs group and the ICBG group was available in 14 studies (1004 participants with BMPs and 766 with ICBG) [15–19,21,22,24,25,27–30,32]. A significant reduction of the reoperation rate was found in subjects receiving lumbar fusion with BMPs (RR = 0.57; 95% CI 0.42–0.77;  $p=0.0002$ ), and no substantial heterogeneity was found ( $I^2=0\%$ ; Fig. 6).

In the secondary outcomes, significant difference was found in the operating time between two groups in 9 trials [14,15,22,23,27,29–32], (MD–0.32; 95% CI–0.55, –0.08;  $P=0.009$ ), it had obviously heterogeneity ( $I^2=79\%$ ; Fig. 7). However, no significant difference was found in the Blood loss between two groups in 8 trials [14,15,22,25,27,30–32], (MD–50.24; 95% CI–117.38, 16.90;  $P=0.14$ ), it also had obviously heterogeneity ( $I^2=77\%$ ; Fig. 8). No significant difference was found in the hospital stay in 7 trials [9,10,15,16,18,20,23] (MD–0.56; 95% CI: –1.12, –0.01;  $P=0.05$ ). It also had obviously heterogeneity ( $I^2=70\%$ ; Fig. 9). Patient satisfaction was available from 4 included studies [15–17,21]. The pooled result showed no significant difference in the BMPs group in comparison to the ICBG group (RR = 1.06; 95% CI 0.86–1.32;  $p=0.58$ ), and a



Figure 2. Risk of bias summary. doi:10.1371/journal.pone.0097049.g002

Published or registered protocols were unavailable for all studies, though we conducted a comprehensive search. In the absence of these, it was difficult for us to decide whether outcomes were measured, but not reported because they were found to be insignificant or unfavorable. Therefore, only eight studies reporting all four primary outcomes (i.e., the solid fusion rate, clinical outcomes, complications, and the reoperation rate) were considered to have fulfilled this criterion [14,15,17,23,25,27,28,32].

The quality of the overall body of evidence for each individual outcome was addressed and summarized through the GRADE system (Table 2). The assessment of the solid fusion rate as a primary outcome was rated as moderate quality, in view of high

**Table 2.** GRADE profile for the quality of evidence related to the assessment for Lumbar Spine using BMPs and ICBG.

Quality assessment		No of patients			Effect	Quality	Importance					
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	BMPs	ICBG	Relative (95% CI)	Absolute		
<b>solid fusion rate (follow-up mean 24 months)</b>												
17	randomised trials	serious <sup>1</sup>	no serious inconsistency	no serious indirectness	no serious imprecision	none	547/610 (89.7%)	410/523 (78.4%)	RR 1.13 (1.05 to 1.23)	102 more per 1000 (from 39 more to 180 more)	⊕⊕⊕○	CRITICAL
<b>overall success of clinical outcomes (follow-up mean 24 months)</b>												
8	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious <sup>3</sup>	reporting bias <sup>2</sup>	339/431 (78.7%)	199/265 (75.1%)	RR 1.04 (0.95 to 1.13)	30 more per 1000 (from 38 fewer to 98 more)	⊕⊕○○	CRITICAL
<b>Complications (follow-up mean 24 months)</b>												
9	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious <sup>3</sup>	none	112/605 (18.5%)	87/444 (19.6%)	RR 0.96 (0.85 to 1.09)	8 fewer per 1000 (from 29 fewer to 18 more)	⊕⊕⊕○	CRITICAL
<b>Reoperation rate (follow-up mean 24 months)</b>												
14	randomised trials	Serious <sup>4</sup>	no serious inconsistency	no serious indirectness	Serious <sup>5</sup>	none	72/1004(7.2%)	100/766(13.1%)	RR 0.57 (0.42 to 0.77)	56 fewer per 1000 (from 30 fewer to 76 fewer)	⊕⊕○○	CRITICAL
<b>operative time (follow-up mean 24 months; Better indicated by lower values)</b>												
9	randomised trials	no serious risk of bias	very serious <sup>6</sup>	no serious indirectness	serious <sup>3</sup>	reporting bias <sup>2</sup>	435	388		44 fewer per 1000 (from 23 fewer to 59 fewer)	⊕○○○	IMPORTANT
<b>Blood loss (follow-up mean 24 months; Better indicated by lower values)</b>												
8	randomised trials	no serious risk of bias	very serious <sup>7</sup>	no serious indirectness	Serious <sup>3</sup>	reporting bias <sup>2</sup>	411	376		MD 50.24 lower (117.38 lower to 16.9 higher)	⊕○○○	IMPORTANT
<b>Hospital stay (follow-up mean 24 months; Better indicated by lower values)</b>												
											⊕○○○	VERY LOW

Table 2. Cont.

Quality assessment		No of patients			Effect	Quality	Importance					
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	BMPs	ICBG	Relative (95% CI)	Absolute		
7	randomised trials	no serious risk of bias	Serious <sup>8</sup>	no serious indirectness	serious <sup>3</sup>	none	377	332		MD 0.56 lower (1.12 to 0.01 lower)	⊕⊕○○	IMPORTANT
<b>Patient satisfaction rate (follow-up mean 24 months)</b>												
4	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	Serious <sup>3</sup>	none	147/186 (79%)	125/165 (75.8%)	RR 1.06 (0.86 to 1.32)	45 more per 1000 (from 106 fewer to 242 more)	⊕⊕⊕○	IMPORTANT
<b>Return-to-work status (follow-up mean 24 months)</b>												
2	randomised trials	serious <sup>9</sup>	serious <sup>10</sup>	no serious indirectness	very serious <sup>11</sup>	none	22/23 (95.7%)	24/27 (88.9%)	RR 1.1 (0.69 to 1.76)	89 more per 1000 (from 276 fewer to 676 more)	⊕⊕○○	IMPORTANT
<b>Work status (follow-up mean 24 months)</b>												
7	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	Serious <sup>3</sup>	none	229/471 (48.6%)	197/416 (47.4%)	RR 1.05 (0.85 to 1.3)	24 more per 1000 (from 71 fewer to 142 more)	⊕⊕⊕○	IMPORTANT
<b>Quality assessment summary</b>												
							42.4%			21 more per 1000 (from 64 fewer to 127 more)	MODERATE	

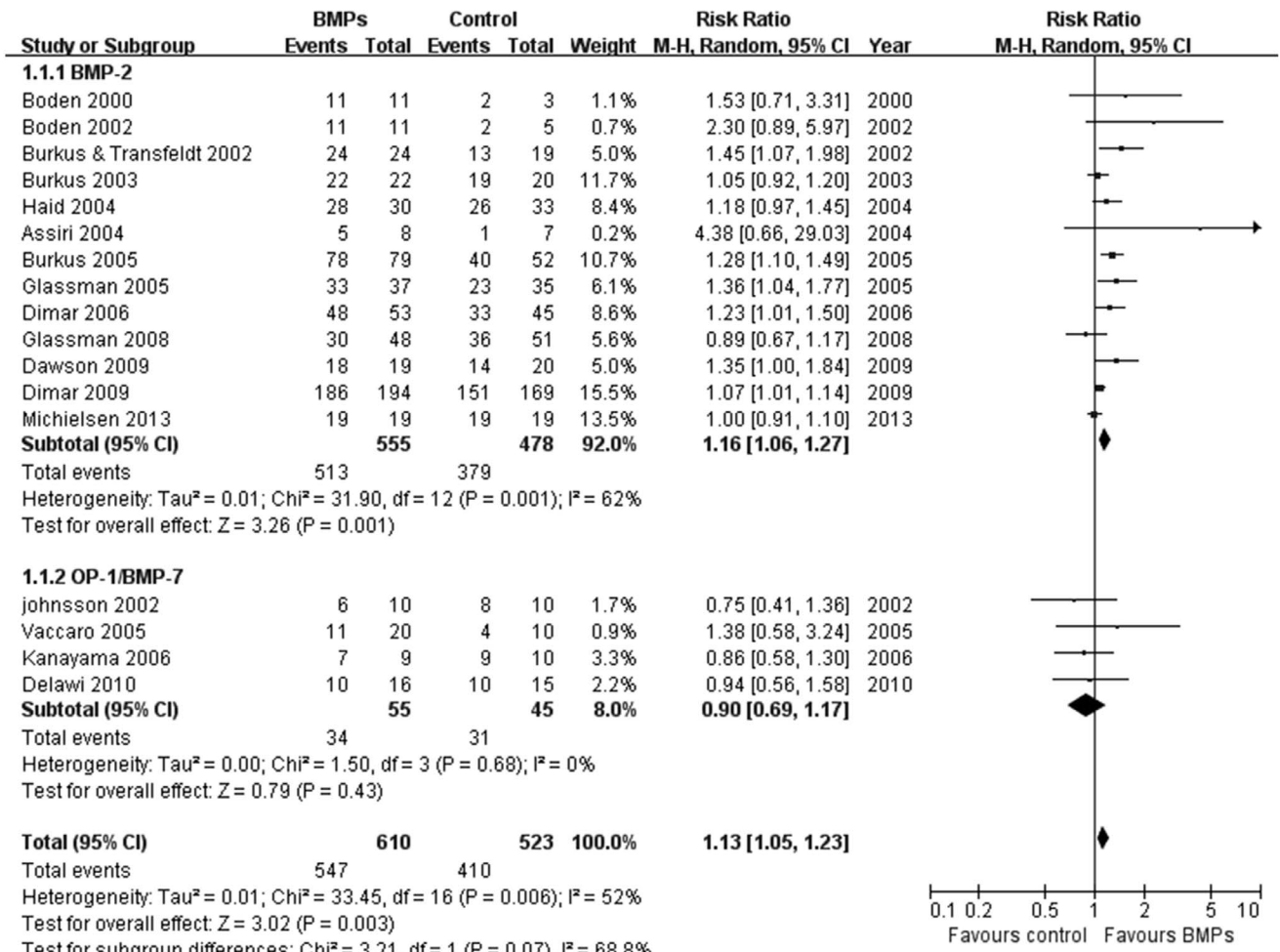
<sup>1</sup>seven studies with a high risk of bias;  
<sup>2</sup>Asymmetry in funnel plot;  
<sup>3</sup>less than 75% of the studies present data that can be included in a meta-analysis;  
<sup>4</sup>five studies with a high risk of bias;  
<sup>5</sup>RR = 0.56;  
<sup>6</sup>I<sup>2</sup> = 79%;  
<sup>7</sup>I<sup>2</sup> = 77%;  
<sup>8</sup>I<sup>2</sup> = 70%;  
<sup>9</sup>one study including more than 50% patients has a high risk of bias;  
<sup>10</sup>I<sup>2</sup> = 70%;  
<sup>11</sup>only two studies including total 50 patients present data that can be included in a meta-analysis.  
 doi:10.1371/journal.pone.0097049.t002

**Table 3.** Clinical relevance.

	Boden 2000	Boden 2002	Boden 2002	Gornet 2002	Transfeldt 2002	Burkus & Burkus 2002	Johnsson 2002	Burkus 2003	Burkus 2004	Assiri 2004	Haid 2004	Burkus 2005	Burkus 2005	Glassman 2005	Vaccaro 2005	Dimar 2006	Kanayama 2006	Glassman 2008	Vaccaro 2008	Dawson 2009	Dimar 2009	Delawi 2010	Michielsen 2013	
1 Are the patients described in detail so that you can decide whether they are comparable to those that you see in your practice?	+	+	+	+	+	+	+	+	?	+	+	+	+	+	?	+	+	+	+	+	+	+	+	+
2 Are the interventions and treatment settings described well enough so that you can provide the same for your patients?	+	+	+	+	+	+	+	+	?	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
3 Were all clinically relevant outcomes measured and reported?	+	+	?	+	+	?	?	-	?	-	-	?	+	+	?	?	+	+	?	+	+	+	+	+
4 Is the size of the effect clinically important?	+	+	+	+	?	?	-	-	?	+	+	+	+	+	-	-	-	-	+	-	-	-	-	+
5 Are the likely treatment benefits worth the potential harms?	+	?	+	+	+	-	+	+	?	+	+	+	+	+	+	?	+	+	+	+	+	+	+	-

doi:10.1371/journal.pone.0097049.t003

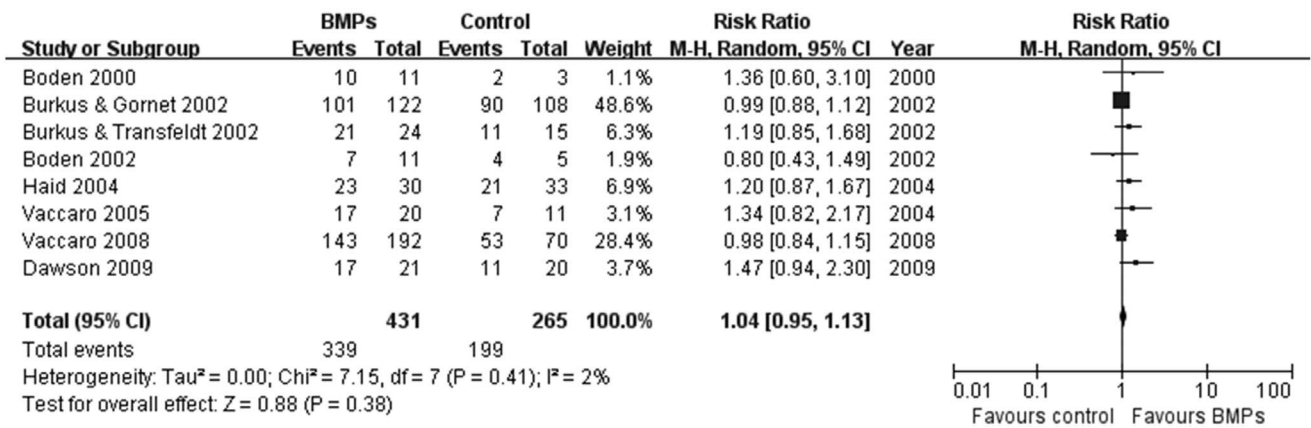




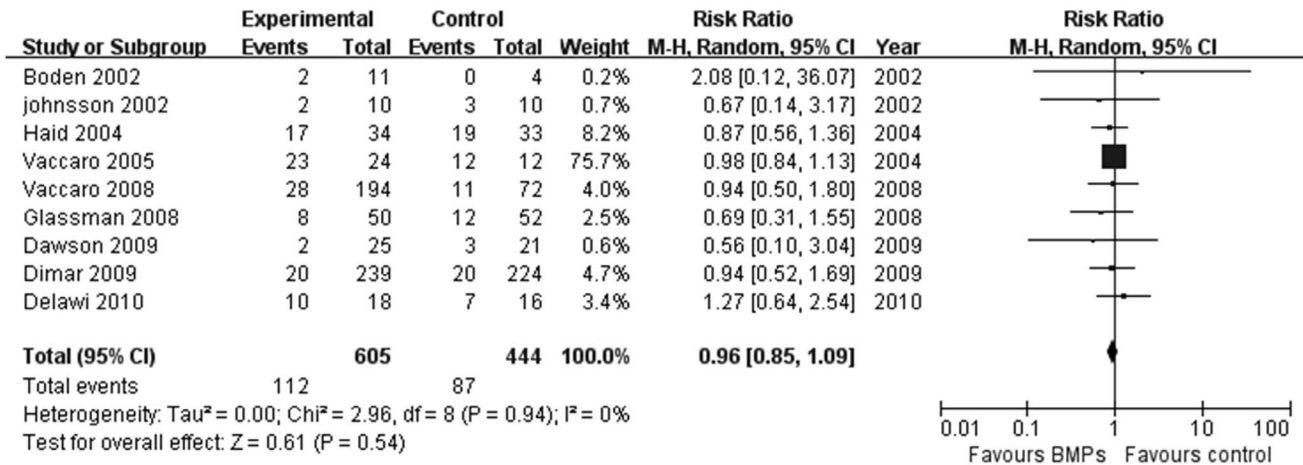
**Figure 3. Forest plot-fusion rate.**  
doi:10.1371/journal.pone.0097049.g003

moderate heterogeneity was found (I<sup>2</sup> = 44%; Fig. 10). The data of patients' work status were available in 7 studies [9,11,12,15,17,18,21] at 24 months follow up. No significant difference was found between two groups (RR = 1.05; 95% CI 0.85–1.30; P = 0.63). There was no significant heterogeneity

between trials (I<sup>2</sup> = 38%; Fig. 11). No significant difference was found about return-to-work status in 2 trials [15,17] (RR 1.10; 95% CI 0.69–1.76; P = 0.68). It also had obviously heterogeneity (I<sup>2</sup> = 70%; Fig. 12).



**Figure 4. Forest plot- overall clinical success.**  
doi:10.1371/journal.pone.0097049.g004



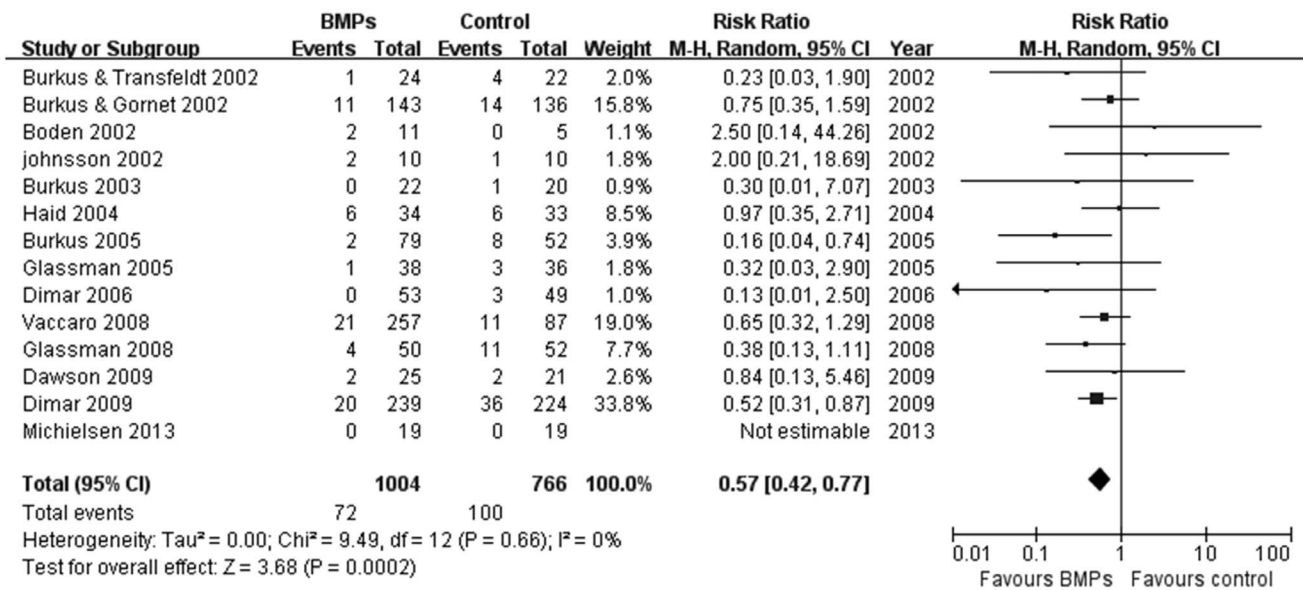
**Figure 5. Forest plot- complications.**  
doi:10.1371/journal.pone.0097049.g005

**Qualitative Data Synthesis**

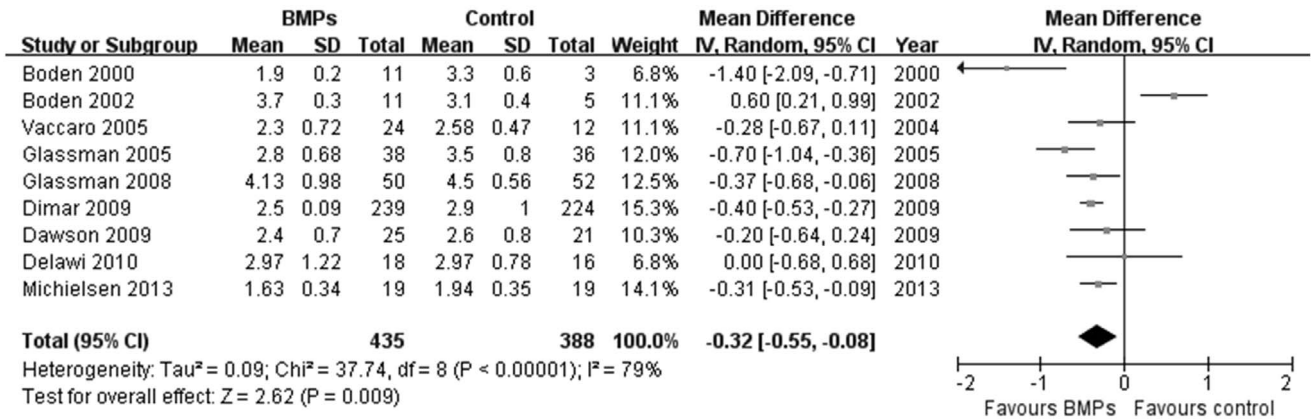
**Donor pain.** Burkus et al [16] found that all the control patients experienced donor site hip pain after surgery. The mean pain score was 12.7 points out of 20 points immediately after surgery, however, at 24 months after surgery pain scores averaged 1.8 points, and 32% patients still experienced pain. In his other study [17], the mean graft-site pain was highest (11.3) after surgery, but it was reduced to 2.2 at 24 months. He also reported 46.5% of the control group patients had persistent pain for 24 months after surgery in the subsequent study in 2005[24]. Haid et al [21] found similar result that the highest levels of pain were noted immediately after surgery with a mean score of 11.6 points, however, at 24 months after surgery, 60% of the control patients still experienced pain, and the graft site pain scores averaged 5.5 points. Dimar et al [25] measured donor site pain utilizing hip pain scores. The mean score after surgery was 11.6, which improved to 7.6 at 24 months after surgery. Vaccaro et al [28] reported 45% of the control group patients had persisted pain for

24 months after surgery, and 35% of the control group patients had persisted mild/moderate pain for 36 months after surgery. Donor site pain was persistent and decreased slowly over time, reported as 1.2 on the VAS (scale of 1–10, 10 being most severe) at 24 months, and 1.1 at 36 months. Dimar et al [29] measured donor site pain using donor-site pain scores. The mean score after discharge was 11.3, which improved to 5.1 at 24 months after surgery, and 60% of the control group patients had persistent pain. Dalewi et al [31] reported that the average donor site pain at 1-year follow-up was graded as 2.7+/-2.8 using the VAS. No complication directly related to the bone graft harvesting procedure occurred.

**Antibody formation.** Six studies assessed antibody responses to BMPs or bovine collagen after surgery. Boden et al [14] did not detect an elevated antibody response to rhBMP-2 in any of the 11 patients, although 3 patients (27%) developed antibodies to bovine type I collagen. No complications were associated with these antibody responses. In the subsequent study, they reported a



**Figure 6. Forest plot- reoperation rate.**  
doi:10.1371/journal.pone.0097049.g006



**Figure 7. Forest plot- operating time.**  
doi:10.1371/journal.pone.0097049.g007

transient antibody response to rhBMP-2 in 1 of 22 patients (4.5%) and 0% (0/4) in the autograft group 3 months after surgery [15]. In Burkus’s study [16], antibodies to rhBMP-2 were evaluated preoperatively and at 3 months after surgery. The results were similar between the rhBMP-2 and control groups. There appeared to be no negative consequence to positive antibody test results. Similarly, 3 months after PLIF with rhBMP-2, Haid et al [21] found that no patients had an elevated antibody response against rhBMP-2, and 3 of 34 patients had developed antibodies against bovine type I collagen. There were no signs of any negative clinical sequelae in patients who tested positive for antibodies against bovine collagen. Burkus et al [24] did not identify an elevated antibody response to rhBMP-2 in any patients, although seven patients (9%) in the study group and four patients (8%) in the control group had an elevated antibody response to bovine collagen. Vaccaro et al [28] found that 25.6% of patients developed neutralizing anti-OP-1 antibodies at any time during follow-up, although there was no association with this neutralizing activity with any clinical outcomes. Further, no neutralizing antibodies were detected in the serum of patients at 24 or 36 month follow-up appointments.

**Sensitivity Analysis**

To evaluate whether the studies rated to be with high risk of bias significantly affected our results, we performed a sensitivity analysis. The methodological quality was assessed using the 12 criteria recommended by the CBRG. A study with a low RoB was defined as one fulfilling six or more of the criteria items. Therefore,

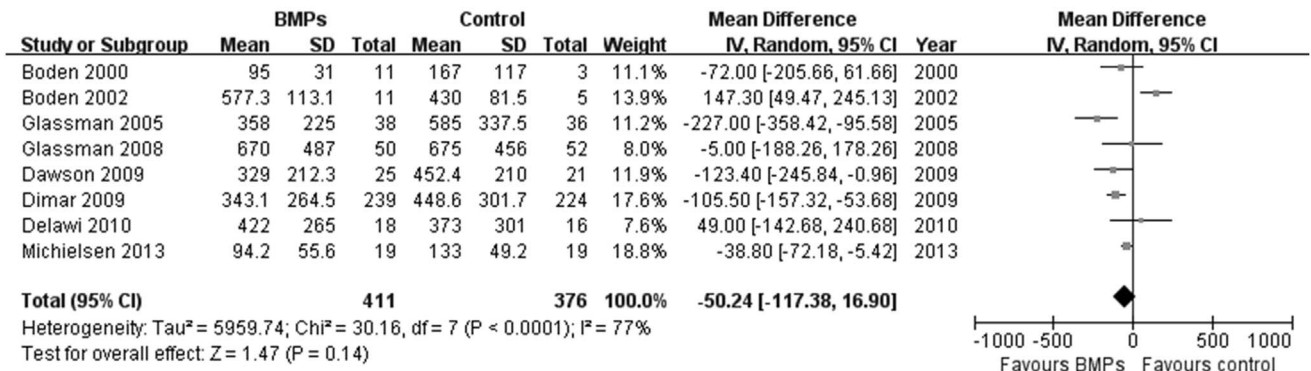
seven studies [19–22,24–26] with a high RoB fulfilling less than six of the 12 criteria items were excluded in sensitive analysis. After excluding these studies, the summary RR of fusion rate at 24 months was 1.09 (95% CI = 0.98–1.21, P = 0.13). These were significantly different from previous results.

**Publication Bias**

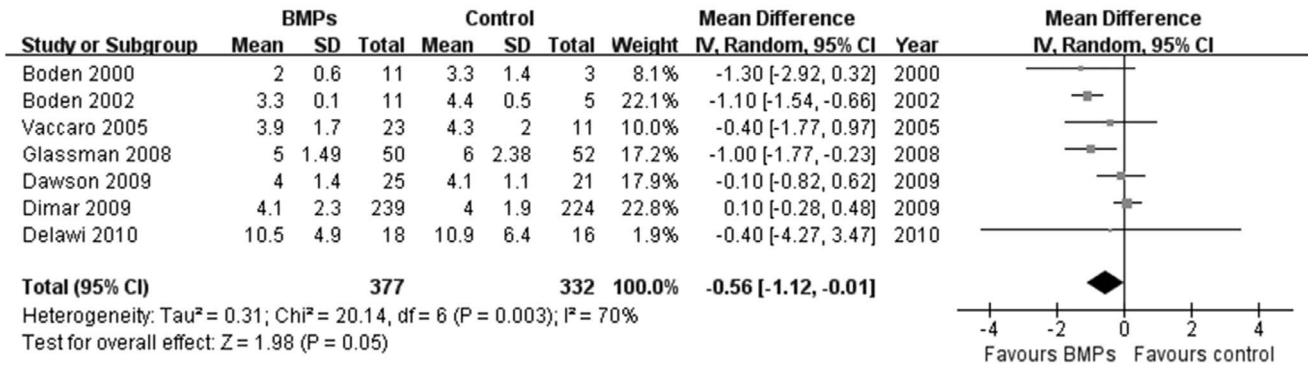
The funnel plot of fusion rate at 24 month is presented in Fig. 13a. No evidences of publication bias were found in both Egger’s test (p = 0.12) and Begg’s test (p = 0.56). However, when we corrected for publication bias using the trim and fill method, the effect of BMPs on fusion rate (RR1.10, 95% CI 1.02–1.19) was not clinically different from the uncorrected result (Fig. 13b).

**Discussion**

The goal of spine surgery for degenerative spinal disease is oftentimes the attainment of solid union of the degenerated and potentially unstable motion segments [33]. Despite the fact that the use of ICBG is the current standard, the morbidity associated with graft harvest has led surgeons to seek viable alternatives [34–38]. BMPs are naturally occurring proteins that stimulate bone healing by a cascade mechanism that results in the differentiation of primitive mesenchymal cells and preosteoblasts into osteoblasts that promote bone formation and, ultimately, healing [39,40]. Currently, two recombinant human BMPs, rhBMP-2 and rhBMP-7, are available for clinical use. These osteoinductive agents have been approved for lumbar fusions either as autologous bone graft



**Figure 8. Forest plot- blood loss.**  
doi:10.1371/journal.pone.0097049.g008



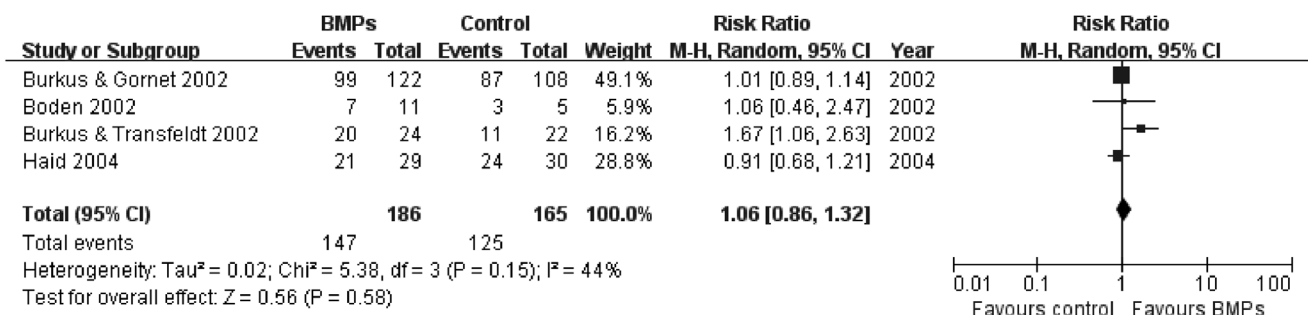
**Figure 9. Forest plot- hospital stay.**  
 doi:10.1371/journal.pone.0097049.g009

enhancers or even autologous bone graft substitutes. However, serious issues and misconceptions regarding the use of osteoinductive bone graft substitutes have recently been outlined [41–43]. So, the purpose of this study is to systematically compare the effectiveness and safety of fusion with BMPs for the treatment of lumbar disease.

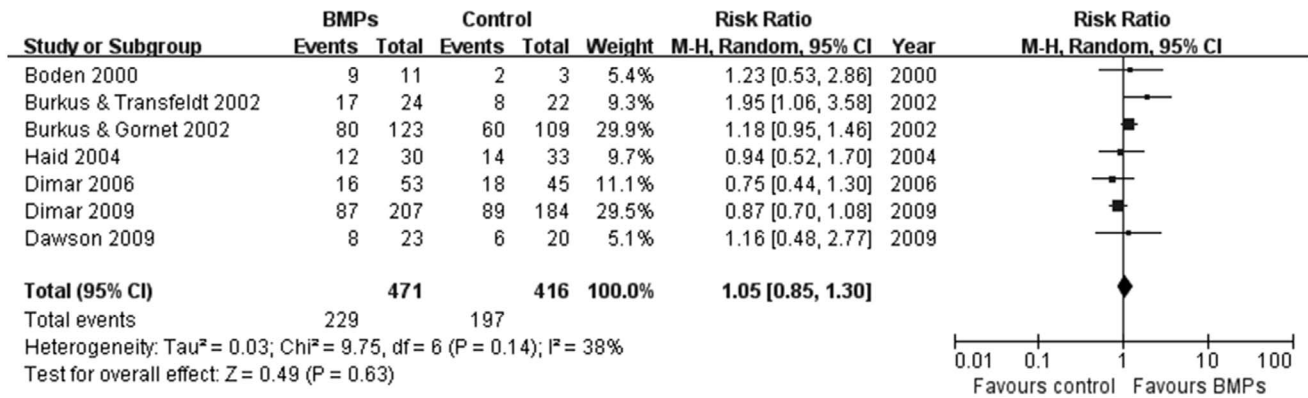
This meta-analysis identified 19 RCTs that compared BMPs and ICBG for lumbar fusion. It revealed that there was significant difference in the solid fusion rate and the reoperation rate. Subgroup analysis of the fusion rate stratified by the two types of BMPs yielded different results. Compared with ICBG, the use of BMP-2 can increase solid fusion rate, by contrast, pooled BMP-7 studies do not have similar effects. However, no significant difference was found in overall success of clinical outcomes and complications. The operating time of BMPs group was shorter than the ICBG group, while the amount of blood loss and hospital stay of BMPs group was not significantly higher than the ICBG group. No significant difference was found in patient satisfaction rate and work status.

Ostensibly, these results are consistent with the previous review. Mussano et al [44] showed that the efficacy of BMPs in vertebral lesions was slightly better than that of standard treatment in terms of producing bone consolidation (radiologic outcome relative risk = 1.07; 95% CI 1.01–1.12), along with functionality and pain (clinical outcome relative risk = 1.08; 95% CI 0.97–1.19). Papakostidis et al [45] evaluated the radiographic and clinical effectiveness of BMPs about lumbar posterolateral fusion. They included seven randomized control trials and one prospective comparative study. Their study found that rhBMP-2 was more efficacious to ICBG in promoting fusion, whereas rhBMP-7 appeared equivalent to ICBG in that respect. Patients treated with BMPs had a shorter hospitalization compared with those that were

treated with ICBG. BMPs appeared more efficient in instrumented than non-instrumented posterolateral fusions. Agarwal et al [46] conducted a systematic review to compare the efficacy and safety of osteoinductive bone graft substitutes using autografts and allografts in lumbar fusion. RhBMP-2 significantly decreased radiographic nonunion compared to ICBG. Trials of rhBMP-2 suggested reductions in the operating time and surgical blood loss, with less effect on the length of hospital stay. There was no difference in radiographic nonunion with the use of rhBMP-7 when compared with ICBG. Neither rhBMP-2 nor rhBMP-7 demonstrated a significant improvement on the ODI when compared with ICBG. Chen et al [47] conducted a systematic review which including ten randomized controlled trials had a conclusion that the use of rhBMP-2 significantly reduced the risk of fusion failure and the rate of reoperation comparing with ICBG. They also find that there was no statistical difference in clinical improvement on the ODI, although a favorable trend in the rhBMP-2 group was found. Donell et al [48] found that the use of BMP-2 was associated with a statistically significantly higher level of spinal fusion than the use of ICBG in patients with single-level DDD. There were no significant differences in the ODI and SF-36 score improvements between BMP-2 and control groups. Adverse events reported were similar between two groups, but one study [21] reported significantly more BMP-2 patients with bone formation outside of the space compared with controls. Recently, serial reports based on Yale University Open Data Access-orchestrated project (YODA) showed different results. Fu et al [49] found that rhBMP-2 has no proven clinical advantage over bone graft and may be associated with important harms, making it difficult to identify clear indications for rhBMP-2. Simmonds et al [50,51] also conducted a individual-participant data meta-analysis (IPDMA). They found that rhBMP-2 increases fusion rates,



**Figure 10. Forest plot- patient satisfaction.**  
 doi:10.1371/journal.pone.0097049.g010



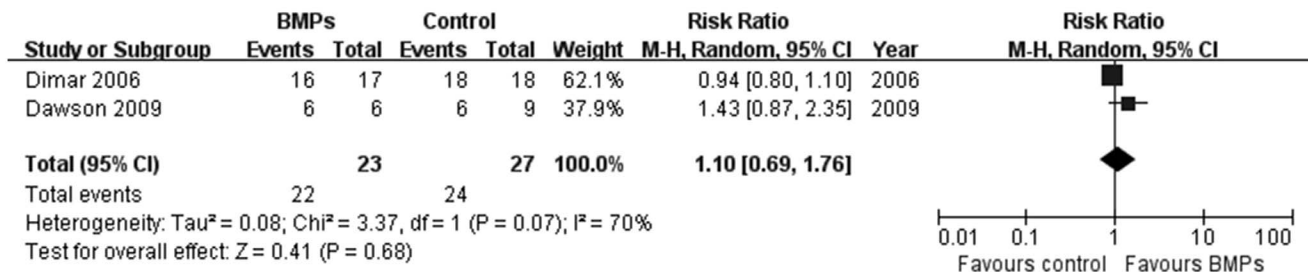
**Figure 11. Forest plot- work status.**  
doi:10.1371/journal.pone.0097049.g011

reduces pain by a clinically insignificant amount, and increases early postsurgical pain compared with ICBG. Evidence of increased cancer incidence is inconclusive.

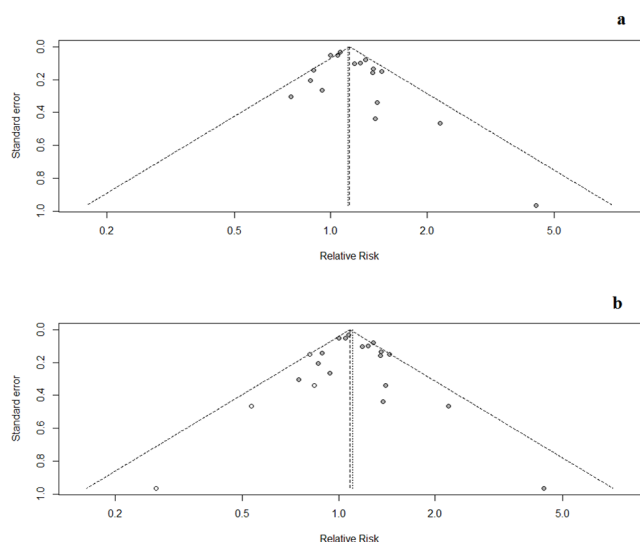
Imaging was used to assess the status of spinal fusion after surgery. However, imaging evaluation is different from the direct operative exploration [52]. Therefore, the fusion rate from imaging evaluation may not equal the actual fusion rate. Furthermore, imaging methods and the fusion standards were variable. Our Meta analysis also included articles utilized plain radiographs and CT-imaging, or surgical exploration as a method of evaluation of fusion status. Thus, the results of sensitivity analysis were significantly different from previous results, probably because of that the validity of the combined results influenced by the potential variability. In our study, there are some excellent fusion results using rhBMP-2 in spinal fusion, but pooled data of fusion using BMP-7 are not inferior to autograft. The unfavorable results may be due to lesser osteoinductive capacity of BMP-7 compared with BMP-2, the lower effective BMP dose, and a different carrier possibly being inferior to the BMP-2 carrier. Although achieving a solid arthrodesis is a primary aim of spinal fusion surgery, the overall goal is to improve quality of life and mobility. We cannot conduct a quantitative synthesis, because of incomplete data of parameters of clinical outcome. However, we described most studies, which reported pain and functional outcome scores between baseline and follow-up. At all follow-up intervals, there were significant improvements in the clinical outcome measures, including the ODI scores, Short Form-36 scores, and back and leg pain scores in both groups, but no significant differences were found between groups. It would seem that the use of BMPs is of no detriment in terms of improvements in functional outcomes.

The purpose of this meta analysis was to evaluate the effectiveness and, more importantly, safety of BMPs compared with ICBG in lumbar fusion. Though some reports lack valid data, a quantitative analysis of complications was conducted, which had no significant difference between BMPs and ICBG group. In a systematic review focusing on the safety of BMP-2, Morz et al [53] determined that multiple complications are associated after the use of rhBMP-2 in both cervical and lumbar spine fusion surgery. There was a mean incidence of 44%, 25%, and 27% of resorption, subsidence, and interbody cage migration reported for lumbar spine interbody fusion surgery although reoperation or long-term detrimental effect was rare. Carragee et al [43] concluded that original industry-sponsored trials underestimated BMP-related adverse events, and they thought the risk of adverse events should be considered in the context of demonstrated benefits. Evidence from YODA serial studies [49–51] also indicated that there appears to have been an increased risk of uncommon and serious complications with the use of BMPs in lumbar fusion. Therefore, in sum, it is difficult for us to determine the nature, range, and frequency of adverse events associated with BMPs.

Our review has limitations. First, the search was restricted to reports of RCTs published in peer-reviewed journals, excluding other sources of biomedical literature, which could have possibly collected more studies related to the topic. In such a case, studies with positive or statistically significant results would be expected to be over represented in our review; such studies are more likely to be published, particularly in the English language. So we used the funnel plot as a tool to investigate how much our results were potentially influenced by publication bias. Second, the validity of our results is limited by the low quality of the studies included, such as double-blinding was unattainable for most of the trials, that may decrease the strength of conclusions drawn from the



**Figure 12. Forest plot- return to work status.**  
doi:10.1371/journal.pone.0097049.g012



**Figure 13. Funnel plot-fusion rate.**  
doi:10.1371/journal.pone.0097049.g013

meta-analysis. Third, there is the potential for bias because device manufacturers sponsored several studies and some authors reported conflicts of interest. However, there were several improvements in this meta-analysis compare with previous systematic reviews. First, this review is the most current report on the topic and includes the recently published trials. It adopted more strict inclusion criteria. Quasi-RCT and non-RCTs were strictly excluded in this study in order to guarantee the reliability of results. Second, we pooled the data of comparable parameters regarding complications to reduce the bias of the descriptive analysis. Third, we also did an additional qualitative data synthesis of donor pain and antibody responses to BMPs. Fourth, the quality of the overall body of evidence for each individual outcome was

## References

- Arrington ED, Smith WJ, Chambers HG, Bucknell AL, Davino NA (1996) Complications of iliac crest bone graft harvesting. *Clin Orthop Relat Res* 329: 300–309. doi:10.1097/00003086-199608000-00037.
- Urist MR (1965) Bone Formation by autoinduction. *Science* 150: 893–899. doi:10.1126/science.150.3698.893.
- Celeste AJ, Iannazzi JA, Taylor RC, Hewick RM, Rosen V, et al. (1990) Identification of transforming growth factor beta family members present in bone-inductive protein purified from bovine bone. *Proc Natl Acad Sci USA* 87: 9843–9847. doi:10.1073/pnas.87.24.9843.
- Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33: 159–174. doi:10.2307/2529310.
- Viera AJ, Garrett JM (2005) Understanding interobserver agreement: the kappa statistic. *Fam Med* 37: 360–363.
- Furlan AD, Pennick V, Bombardier C, van Tulder M (2009) 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. *Spine*. doi:10.1097/BRS.0b013e3181b1c99f.
- Higgins JPT, Green S (2008) *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.0.1 [updated September 2008]. The Cochrane Collaboration.
- DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. *Control Clin Trials* 7: 177–188. doi:10.1016/0197-2456(86)90046-2.
- Egger M, Smith GD, Schneider M, Minder C (2000) Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315: 629–634. doi:10.1136/bmj.315.7109.629.
- Begg CB, Mazumdar M (1994) Operating characteristics of a rank correlation test for publication bias. *Biometrics* 50: 1088–1101. doi:10.2307/2533446.
- Egger M, Smith GD (1998) Bias in location and selection of studies. *BMJ* 316: 61–66. doi:10.1136/bmj.316.7124.61.
- Duval S, Tweedie R (2000) Trim and fill: a simple funnel-plot based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 56: 455–463. doi:10.1111/j.0006-341X.2000.00455.x.
- Sutton AJ, Duval SJ, Tweedie RL, Abrams KR, Jones DR (2000) Empirical assessment of effect of publication bias on meta-analyses. *BMJ* 320: 1574–1577. doi:10.1136/bmj.320.7249.1574.
- 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* 25: 376–381. doi:10.1097/00007632-200002010-00020.
- 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* 27: 2662–2673. doi:10.1097/00007632-200212010-00005.
- 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: 337–349. doi:10.1097/00024720-200210000-00001.
- 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* 27: 2396–2408. doi:10.1097/00007632-200211010-00015.
- Johnsson R, Strömqvist B, Aspenberg P (2002) Randomized radiostereometric study comparing osteogenic protein-1 (BMP-7) and autograft bone in human noninstrumented posterolateral lumbar fusion: 2002 Volvo Award in clinical studies. *Spine* 27: 2654–2661. doi:10.1097/00007632-200212010-00004.
- Burkus JK, Dorchak JD, Sanders DL (2003) Radiographic assessment of interbody fusion using recombinant human bone morphogenetic protein type 2. *Spine* 28: 372–377. doi:10.1097/01.BRS.0000048469.45035.B9.
- 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: 527–538. doi:10.1016/j.spinee.2004.03.025.
- Assiri I, du Plessis S, Hurlbert J, Hu R, Salo P, et al. (2004) A prospective randomized clinical study comparing instrumented lumbar fusion rates of Recombinant Human Bone Morphogenetic Protein-2 (rhBMP-2) with autogenous

addressed and summarized through the GRADE system, that provided a better guideline for the clinical practice.

## Conclusion

In summary, our review adds to the evidence concerning the use of BMPs for lumbar fusion. Various RCT studies conclude that the use of BMPs can increase the fusion rate slightly, while decrease the reoperation rate and operating time. There was no significant difference in the overall success of clinical outcome, the complication rate, the amount of blood loss and hospital stay between the two groups. The use of BMPs prevents graft site related adverse effects. No complications were associated with antibody responses. From the limited evidence, BMPs does not show significant superiority for the treatment of LDD compared with ICBG. To assess the effectiveness and safety of lumbar fusion with BMPs, more high-quality RCTs with long term outcomes are needed.

## Supporting Information

### Checklist S1 PRISMA 2009 Checklist.

(DOC)

### Search strategies S1

(DOC)

### List S1 List of included-excluded studies.

(DOC)

## Author Contributions

Conceived and designed the experiments: HZ FW LD ZZ YZ DS XF JA. Analyzed the data: HZ FW LD ZZ XF YZ JA. Wrote the paper: HZ FW LD ZZ YZ DS XF JA. Screened the literature search results and identified articles for retrieval, reviewed full papers for inclusion, and extracted data for meta-analysis: HZ FW DS XF. Carried out comparisons: HZ FW LD ZZ XF YZ JA.

- iliac crest bone graft in patients with symptomatic degenerative disc disease. *Canadian Journal of Surgery (Suppl 4)*: 7–8.
22. Glassman SD, Dimar JR, Carreon LY, Campbell MJ, Puno RM, et al. (2005) Initial fusion rates with recombinant human bone morphogenetic protein-2/compression resistant matrix and a hydroxyapatite and tricalcium phosphate/collagen carrier in posterolateral spinal fusion. *Spine* 30: 1694–1698. doi:10.1097/01.brs.0000172157.39513.80.
  23. Vaccaro AR, Anderson DG, Patel T, Fischgrund J, Truumees E, et al. (2005) Comparison of OP-1 Putty (rhBMP-7) to iliac crest autograft for posterolateral lumbar arthrodesis: a minimum 2-year follow-up pilot study. *Spine* 30: 2709–2716. doi:10.1097/01.brs.0000190812.08447.ba.
  24. Burkus JK, Sandhu HS, Gornet MF, Lonely 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: 1205–1212. doi:10.2106/JBJS.D.02532.
  25. Dimar JR, Glassman SD, Burkus KJ, Carreon LY (2006) Clinical outcomes and fusion success at 2 years of single-level instrumented posterolateral fusions with recombinant human bone morphogenetic protein-2/compression resistant matrix versus iliac crest bone graft. *Spine* 31: 2534–2539. doi:10.1097/01.brs.0000240715.78657.81.
  26. Kanayama M, Hashimoto T, Shigenobu K, Yamane S, Bauer TW, et al. (2006) A prospective randomized study of posterolateral lumbar fusion using osteogenic protein-1 (OP-1) versus local autograft with ceramic bone substitute: emphasis of surgical exploration and histologic assessment. *Spine* 31: 1067–1074. doi:10.1097/01.brs.0000216444.01888.21.
  27. Glassman SD, Carreon LY, Djurasovic M, Campbell MJ, Puno RM, et al. (2008) RhBMP-2 versus iliac crest bone graft for lumbar spine fusion: a randomized, controlled trial in patients over 60 years of age. *Spine* 33: 2843–2849. doi:10.1097/BRS.0b013e318190705d.
  28. Vaccaro AR, Lawrence JP, Patel T, Katz LD, Anderson DG, et al. (2008) The safety and efficacy of OP-1 (rhBMP-7) as a replacement for iliac crest autograft in posterolateral lumbar arthrodesis: a long-term (>4 years) pivotal study. *Spine* 33: 2850–2862. doi:10.1097/BRS.0b013e3181818a314d.
  29. Dimar JR 2nd, Glassman SD, Burkus JK, Pryor PW, Hardacker JW, et al. (2009) Clinical and radiographic analysis of an optimized rhBMP-2 formulation as an autograft replacement in posterolateral lumbar spine arthrodesis. *Journal Bone Joint Surg Am* 91: 1377–1386. doi:10.2106/JBJS.H.00200.
  30. Dawson E, Bac 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: 1604–1613. doi:10.2106/JBJS.G.01157.
  31. Delawi D, Dhert WJ, Rillardon L, Gay E, Prestamburgo D, et al. (2010) A prospective, randomized, controlled, multicenter study of osteogenic protein-1 in instrumented posterolateral fusions: report on safety and feasibility. *Spine* 35: 1185–1191. doi:10.1097/BRS.0b013e3181d3cf28.
  32. Michielsen J, Sys J, Rigaux A, Bertrand C (2013) The effect of recombinant human bone morphogenetic protein-2 in single-level posteriolateral interbody arthrodesis. *J Bone Joint Surg Am* 95: 873–880. doi:10.2106/JBJS.L.00137.
  33. An H, Boden SD, Kang J, Sandhu HS, Abdu W, et al. (2003) Summary statement: emerging techniques for treatment of degenerative lumbar disc disease. *Spine* 28: S24–S25. doi:10.1097/01.BRS.0000076894.33269.19.
  34. Younger EM, Chapman MW (1989) Morbidity at bone graft donor sites. *J Orthop Trauma* 3: 192–195. doi:10.1097/00005131-198909000-00002.
  35. Banwart JC, Asher MA, Hassancin RS (1995) Iliac crest bone graft harvest donor site morbidity: a statistical evaluation. *Spine* 20: 1055–1060. doi:10.1097/00007632-199505000-00012.
  36. Hutchinson MR, Dall BE (1994) Midline fascial splitting approach to the iliac crest for bone graft: a new approach. *Spine* 19: 62–66. doi:10.1097/00007632-199401000-00013.
  37. Ahlmann E, Patzakis M, Roidis N, Shepherd L, Holtom P (2002) Comparison of anterior and posterior iliac crest bone grafts in terms of harvest-site morbidity and functional outcomes. *J Bone Joint Surg Am* 84: 716–720.
  38. Sen MK, Miclau T (2007) Autologous iliac crest bone graft: should it still be the gold standard for treating nonunions? *Injury* 38 Suppl 1: S75–80. doi:10.1016/j.injury.2007.02.012.
  39. Reddi AH (1998) Role of morphogenetic proteins in skeletal tissue engineering and regeneration. *Nat Biotechnol* 16: 247–252. doi:10.1038/nbt0398-247.
  40. Dimitriou R, Tsiridis E, Giannoudis PV (2005) Current concepts of molecular aspects of bone healing. *Injury* 36: 1392–1404. doi:10.1016/j.injury.2005.07.019.
  41. Poynton AR, Lane JM (2002) Safety profile for the clinical use of bone morphogenetic proteins in the spine. *Spine* 27: S40–48. doi:10.1097/00007632-200208151-00010.
  42. Benglis D, Wang MY, Levi AD (2008) A comprehensive review of the safety profile of bone morphogenetic protein in spine surgery. *Neurosurgery* 62: ONS423–431; discussion ONS431. doi:10.1227/01.neu.0000326030.24220.d8.
  43. Carragee EJ, Hurwitz EL, Weiner BK (2011) A critical review of recombinant human bone morphogenetic protein-2 trials in spinal surgery: emerging safety concerns and lessons learned. *Spine J* 11: 471–491. doi:10.1016/j.spinee.2011.04.023.
  44. Mussano F, Ciccone G, Ceccarelli M, Baldi I, Bassi F (2007) Bone morphogenetic proteins and bone defects: a systematic review. *Spine* 32: 824–830. doi:10.1097/01.brs.0000259227.51180.ca.
  45. Papakostidis C, Kontakis G, Bhandari M, Giannoudis PV (2008) Efficacy of autologous iliac crest bone graft and bone morphogenetic proteins for posterolateral fusion of lumbar spine: a meta-analysis of the results. *Spine* 33: E680–692. doi:10.1097/BRS.0b013e31818184eca.
  46. Agarwal R, Williams K, Umscheid CA, Welch WC (2009) Osteoinductive bone graft substitutes for lumbar fusion: a systematic review. *J Neurosurg Spine* 11: 729–740. doi:10.3171/2009.6.SPINE08669.
  47. Chen Z, Ba G, Shen T, Fu Q (2012) 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* 132: 1725–1740. doi:10.1007/s00402-012-1607-3.
  48. Donell S, Garrison K, Shemilt I, Song F (2011) Bone Morphogenetic Protein-2 for Spinal Fusion: Meta-analysis of Randomised Controlled Trials. UEA Website. Available: <http://www.uea.ac.uk/~w158/BMP-2%20for%20SF%20-%20Meta-analysis%20of%20RCTs.pdf>. Accessed Dec 12, 2013.
  49. Fu R, Selph S, McDonagh M, Peterson K, Tiwari A, et al. (2013) Effectiveness and harms of recombinant human bone morphogenetic protein-2 in spine fusion: a systematic review and meta-analysis. *Ann Intern Med* 158: 890–902. doi:10.7326/0003-4819-158-12-201306180-00006.
  50. Simmonds MC, Brown JV, Heirs MK, Higgins JP, Mannion RJ, et al. (2013) Safety and effectiveness of recombinant human bone morphogenetic protein-2 for spinal fusion: a meta-analysis of individual-participant data. *Ann Intern Med* 158: 877–889. doi:10.7326/0003-4819-158-12-201306180-00005.
  51. Yale School of Medicine. Available: [http://medicine.yale.edu/core/projects/yodap/datasharing/medtronic/463\\_158787\\_YorkrhBMP2FinalReport.pdf](http://medicine.yale.edu/core/projects/yodap/datasharing/medtronic/463_158787_YorkrhBMP2FinalReport.pdf). Accessed Dec 12, 2013.
  52. Blumenthal SL, Gill K (1993) Can lumbar spine radiographs accurately determine fusion in postoperative patients? Correlation of routine radiographs with a second surgical look at lumbar fusions. *Spine* 18: 1186–1189. doi:10.1097/00007632-199307000-00010.
  53. Mroz TE, Wang JC, Hashimoto R, Norvell DC (2010) Complications related to osteobiologics use in spine surgery: a systematic review. *Spine* 35: S86–104. doi:10.1097/BRS.0b013e3181d81cf2.