Spine

SPINE Volume 45, Number 12, pp E729–E741

LITERATURE REVIEW

OPEN

Comparative Clinical Effectiveness and Safety of Bone Morphogenetic Protein Versus Autologous Iliac Crest Bone Graft in Lumbar Fusion

A Meta-analysis and Systematic Review

Shichao Liu, MD, Yinqing Wang, MD, Zeyan Liang, MD, Maochao Zhou, MD, and Chunmei Chen, MD, PhD

Study Design. This is a systematic literature review and meta-analysis.

Objective. We aimed to evaluate the efficacy and safety of recombinant human bone morphogenetic protein (RhBMP) and autologous iliac crest bone graft (ICBG) in lumbar fusion.

Summary of Background Data. RhBMP has been emphasized in lumbar fusion due to high fusion success rate. However, ICBG remains the criterion standard graft approach for lumbar fusion. The safety and effectiveness of rhBMP are controversial.

Methods. Prospective randomized controlled trials were searched from PubMed, EMBASE, and Cochrane Central Register of Controlled Trails by using Medical Subject Headings terms ''bone morphogenetic protein," ''bone transplantation," and ''spinal fusion." Two independent investigators screened eligible studies, assessed the bias of original articles, extracted data including fusion success, Oswestry disability index improvement, improved short form 36 questionnaire scores, adverse events and re-operation, and a subgroup analysis. The GRADE approach was used to grade quality of evidence.

Results. Twenty randomized controlled trials (2185 patients) met the inclusion criteria. There were higher fusion success rate (odds ratio [OR] 3.79, 95% confidence interval [CI] 1.88–7.63, $P = 0.0002$), better improvement of Oswestry Disability Index (mean difference 1.54, 95% CI 0.18–2.89, $P = 0.03$), and lower re-operation rate (OR 0.59, 95% CI 0.43–0.80, $P = 0.0007$) in rhBMP group. Heterogeneity was obvious in fusion success rate ($I^2 = 58\%$); hence, a subgroup analysis, based on protein type (rhBMP-2 or rhBMP-7), was performed, which suggested that only rhBMP-2 was better than ICBG for lumbar fusion. There was no difference in the incidence of adverse events between rhBMP and ICBG (OR 0.91, 95% CI 0.70–1.18, $P = 0.47$).

Conclusion. In lumbar fusion, rhBMP-2 exhibited a higher fusion success rate and reduced the risk of re-operation. No difference in complication rate is between rhBMP (rhBMP-2 and rhBMP-7) and ICBG. We suggest rhBMP especially rhBMP-2 as an effective substitute for ICBG for lumbar fusion.

Key words: adverse events, autologous iliac crest bone graft, bone morphogenetic protein, fusion success, lumbar fusion, meta-analysis, randomized controlled trial, reoperation.

Level of Evidence: 1

Spine 2020;45:E729–E741

Chronic low back pain and leg pain are commonly caused or influenced by lumbar degenerative diseases, which in include lumbar spinal stenosis, lumbar spondylolisthesis, and lumbar disc herniation.[1] Some patients' condition and symptoms may be controlled with conservative therapies. However, if symptoms or imaging indicate vertebral or spinal instability, lumbar fusion is recommended to re-establish stability.[2] Surgical options for interbody fusion include anterior lumbar interbody fusion (ALIF), posterior lumbar interbody fusion (PLIF), posterolateral lumbar fusion (PLF), or transforaminal lumbar interbody fusion.[3]

To improve success rate of intervertebral fusion, bone graft is considered necessary. For decades, autologous iliac bone graft (ICBG) has been recognized as the criterion standard for lumbar fusion.[4] However, it has disadvantages including an elevated rate of bone nonunion, donor site complications, and relatively insufficient grafted bone in multisegment fusion.[5,6] As a result, various bone graft substitutes have been developed. One of the most widely

From the Department of Neurosurgery, Fujian Medical University Union Hospital, Fuzhou, Fujian, China.

Acknowledgment date: September 25, 2019. First revision date: October 21, 2019. Second revision date: November 5, 2019. Acceptance date: November 20, 2019.

The manuscript submitted does not contain information about medical device(s)/drug(s).

Fujian Provincial Medical Innovation Project of China (Grant No. 2017-CX-27) funds were received in support of this work.

No relevant financial activities outside the submitted work.

Address correspondence and reprint requests to Chunmei Chen, MD, PhD, Department of Neurosurgery, Fujian Medical University Union Hospital, No. 29, Xinquan Road, Fuzhou, Fujian 350001, China;
E-mail: cmchen2009@sina.com.

DOI: 10.1097/BRS.0000000000003372

used is bone morphogenetic protein (BMP), first reported by Urist *et al*,[7] which can induce osteogenesis. Recombinant human bone morphogenetic protein (rhBMP) was developed on a large scale in mid-1990s using human recombinant genetic technology due to the limited yield of natural extracted and purified BMP.[8] The US Food and Drug Administration approved clinical use of two rhBMP: rhBMP-2 and rhBMP-7 as alternatives to ICBG.[9]

Since Boden *et al*[10] conducted the first RCT comparing rhBMP-2 with ICBG, >20 RCTs have been performed. RhBMP has demonstrated advantages in most studies, including higher fusion rate and shorter operation time than ICBG. However, the safety of rhBMP has been questioned due to reports linking the therapy with serious complications.[11–14] Noshchenko *et al* and Zhang *et al* suggested that rhBMP could be a good alternative approach to ICBG.[15,16] However, Ye *et al*[17] indicated that using rhBMP-7 appeared to yield lower fusion rate in instrumented posterolateral fusion patients. A degree of uncertainty remains concerning efficacy and safety of rhBMP in lumbar fusion. In light of this, the purpose of present study is to further update this topic and attempt to clarify rhBMP safety and effects in lumbar fusion.

MATERIALS AND METHODS

Search Strategy

We comprehensively searched PubMed, EMBASE, and Cochrane Central Register of Controlled Trails databases for RCTs, through using MeSH terms "bone morphogenetic protein," "bone transplantation," "bone graft," and "spinal fusion." The retrieved results were last updated on May 29, 2019. To conduct a thorough search of all relevant literature, two independent investigators screened eligible studies. When consensus could not be reached, a third reviewer was consulted to resolve the disagreement.

Study Selection Criteria

Inclusion criteria for articles were: age 18 to 80 years, suffering from lumbar degenerative diseases requiring lumbar fusion, and RCT comparing rhBMP with ICBG.

Studies were excluded if patients presented with spinal deformities, fractures, tumors or infections, cases demonstrated spondylolisthesis classified higher than Meyerding Grade II, follow-up was <12 months, and there were incomplete follow-up data.

Data Extraction

Two investigators extracted data from forms containing relevant patient information. The data included study design, patient characteristics, sample size, intervention details, follow-up rate and time, outcomes. The primary outcomes included fusion success, improvement on the Oswestry disability index (ODI),[18] improvement on short form 36 (SF-36),[19] improvement on the Numeric Rating Scale (NRS) for back pain and leg pain,[20] adverse events, and reoperation. Fusion success was defined as an absence of radiolucent lines covering >50% of either implant, translation of 3 mm, and angulation <5° on flexion–extension radiographs.[21] SF-36 includes Physical Component Summary (PCS) and Mental Component Summary (MCS). Improvement value was defined as absolute difference between preoperative and postoperative outcomes. Secondary outcomes included operation time, intraoperative blood loss, and duration of hospital stay.

Quality Assessment

Two investigators evaluated bias risk using the 12 criteria recommended by the Cochrane Back Review Group.[22] The items were scored as "low risk," "high risk," or "unclear." If at least six of the criteria passed without serious potential flaws, items were considered to have an overall "low risk of bias." Our study utilized the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) group criteria to describe both quality of evidence and strength of recommendations.[23] The quality of evidence was classified as very low, low, moderate, or high.

Statistical Analysis

The results were expressed in terms of an odds ratio (OR) and a 95% confidence interval (95% CI) for dichotomous outcomes, and in terms of mean difference (MD) and 95% CI for continuous outcomes. The I^2 test was used to evaluate the heterogeneity of statistical results. When the I^2 value was <50%, statistical results were considered to have no heterogeneity. Also, a fixed-effect model was used. Otherwise, a random-effect model was used. Funnel plots were used to explore potential publication bias. Meta-analysis was performed by Review Manager Software (RevMan, Version 5.3).

RESULTS

Search Results and Study Characteristics

A total of 1118 related studies were initially identified from PubMed (n = 482), EMBASE (n = 599), and CENTRAL (n = 37) databases. Only 24 RCTs were included.[10,14,24–45] One of these RCTs included follow-up of <12 months,[24] two were duplicate reports of the same set of patients,[34,40] and one reported missing rate >15%.[36] Ultimately, 20 RCTs with 2185 patients were included in meta-analysis. The process of identifying related reports is presented in Figure 1. The extracted data of the characteristics of 20 studies were recorded (Table 1).[10,14,25–33,35,37–39,41-45]

Primary Outcomes

Fusion Success

Thirteen studies met the designated fusion success criteria.[10,26,29,30,32,33,37–39,41,42,44,45] Significant differences were found between rhBMP and ICBG (odds ratio [OR] 3.79, 95% confidence interval [CI] 1.88–7.63, Figure 2). Subgroup analysis based on protein type was performed. The "rhBMP-2 group" comprised 10 articles including

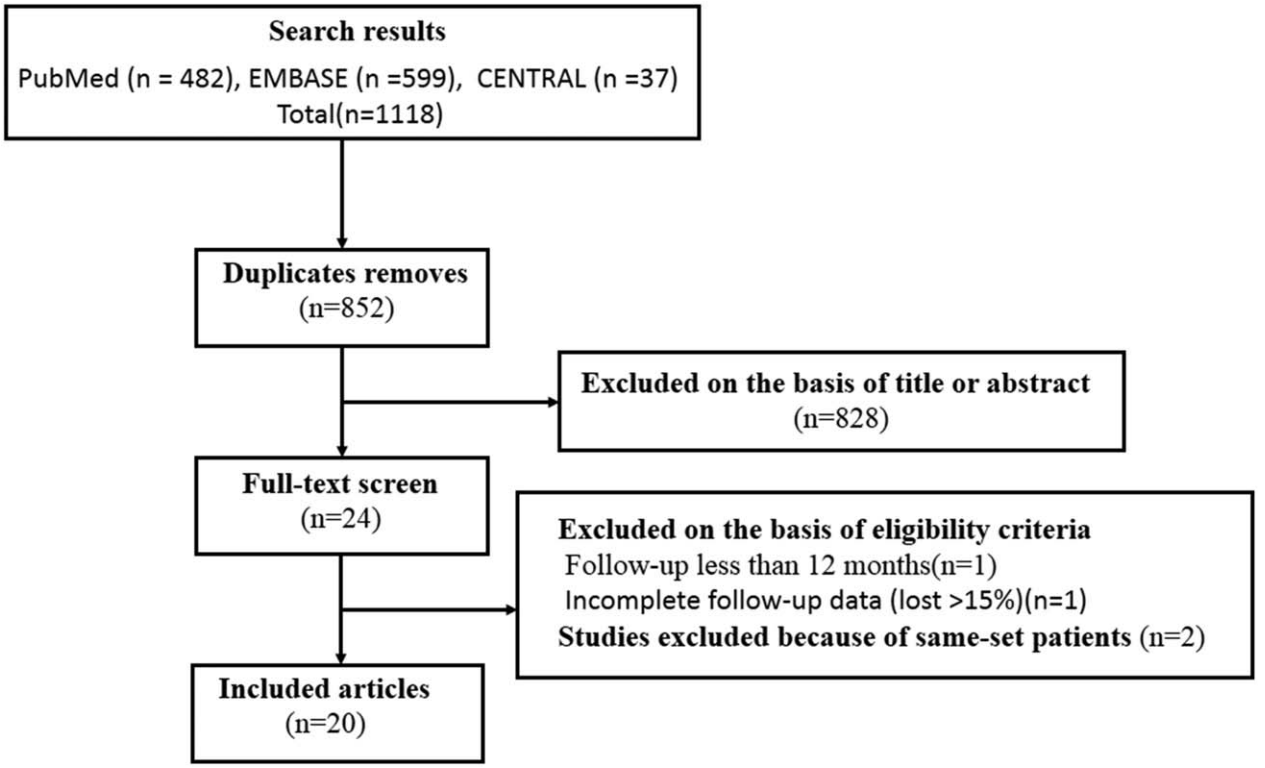

Figure 1. Inclusion and exclusion flow chart.

1141 patients[10,26,29,30,37,39,41,42,44,45] and the ''rhBMP-7 group'' comprised three articles including 245 patients.[32,33,38] In rhBMP-2 group, significant differences were observed between rhBMP-2 and ICBG (OR 5.57, 95% CI 2.95–10.52). In rhBMP-7 group, no significant difference was observed between rhBMP-7 and ICBG with regard to fusion success (OR 0.94, 95% CI 0.49–1.81).

Improvement on ODI

Eleven eligible studies including 1252 patients reported the improvement on ODI.[10,25–32,42,45] Meta-analysis discovered significant difference between rhBMP and ICBG (MD 1.54, 95% CI 0.18–2.89, Figure 3) and there was significant heterogeneity between trials. Subgroup analysis based on protein type was performed. Significant difference was observed between rhBMP-2 and ICBG (MD 1.94, 95% CI 0.28–3.61). No significant difference was observed between rhBMP-7 and ICBG (MD 0.19, 95% CI −2.18 to 2.56).

Improvement on SF-36 PCS

Details regarding improvement on SF-36 PCS on rhBMP-2 were available for four studies including 564 patients.[27,29,31,42] No significant difference was observed between rhBMP-2 and ICBG (MD 1.16, 95% CI −0.84 to 3.16). With regard to rhBMP-7, there were two studies[32,38] that reported data, but omitted the means and standard deviations. As a result, meta-analysis was not performed between rhBMP-7 and ICBG.

Improvement on NRS back Pain and NRS Leg Pain

Relevant improvement on NRS back pain and NRS leg pain data on rhBMP-2 were extracted from five eligible articles containing 755 patients.[26,29,31,42,45] No significant difference in NRS back pain was identified between rhBMP-2 and ICBG (MD −0.05, 95% CI −1.15 to 1.06). With respect to NRS leg pain, a significant difference was observed between rhBMP-2 and ICBG (MD −0.84, 95% CI 0.02–1.65). Due to studies[25,28,32,33,38,43] lacking corresponding data on rhBMP-7, meta-analysis was not performed between rhBMP-7 and ICBG.

Reoperation

Reoperation information was available from 17 studies. Thirteen[14,26,27,29–31,35,37,39,41,42,44,45] of these studies including 1594 patients were about rhBMP-2, and four[25,32,38,43] more including 519 patients were about rhBMP-7. The combined result showed a significantly lower rate of reoperation in rhBMP compared to ICBG (OR 0.59, 95% CI 0.43–0.80, Figure 4). Subgroup analysis revealed a significant difference between rhBMP-2 and ICBG (OR 0.48, 95% CI 0.33–0.68), with no significant difference observed between rhBMP-7 compared with ICBG (OR 1.18, 95% CI 0.62–2.23).

Adverse Events

Fourteen studies provided specific data regarding adverse events. Nine[10,26,27,29–31,39,42,44] of these studies were about rhBMP-2, and five[25,28,32,38,43] more were about rhBMP-7.

TABLE 1. Characteristics of Included Studies.

Study	Study Type	Preoperative Diagnosis	Surgical Intervention	Comparisons	Sample Size		Mean Age, y		BMP Dose, mg	Follow-up, mo	Outcomes
					BMP	ICBG	BMP	ICBG			
Boden et al, 2000[10]	RCT	Lumbar DDD, grade I spondylolisthesis	ALIF	rhBMP-2/ACS *vs.* ICBG	11	3	42.5	40.2	NR	24	FS, ODI, SF-36, AD, OT, LOS, BL
Boden et al, 2002[42]	RCT	Lumbar DDD, grade I spondylolisthesis	ALIF	rhBMP-2/BCP *vs.* ICBG	11	5	57.6	52.9	20	17	FS, ODI, SF-36, reoperation, NRS, AD, OT, LOS, BL
Burkus et al, 2002[44]	RCT	Lumbar DDD	ALIF	rhBMP-2/ACS *vs.* ICBG	143	136	43.3	42.3	4.2–8.4	24	FS, ODI, SF-36, reoperation, neurologic status, NRS, OT, LOS, BL
Burkus et al , 2002[20]	RCT	Lumbar DDD	ALIF	rhBMP-2/ACS *vs.* ICBG	24	22	41.5	45.6	12–18	24	FS, ODI, SF-36, reoperation, AD, neurologic status, NRS, OT, LOS, BL
Johnsson et al, 2002[43]	RCT	L5 spondylolysis and vertebral slip ≤50%	PLIF	rhBMP-7/Type 1 bone collagen *vs.* ICBG	10	10	42.9	40.4	7	12	FS, reoperation, AD
Burkus et al, 2003[41]	RCT	Degenerative lumbar spondylosis	ALIF	rhBMP-2/ACS *vs.* ICBG	22	20	41.7	44.2	4.2–8.4	24	FS, reoperation
Haid et al, 2004[39]	RCT	Lumbar DDD, grade I spondylolisthesis	PLIF	rhBMP-2/ACS *vs.* ICBG	34	33	46.3	46.1	4–8	24	FS, ODI, SF-36, reoperation, AD, neurologic status, NRS, OT, LOS, BL
Glassman et al, 2005[35]	RCT	Lumbar DDD, grade I spondylolisthesis	PLF	rhBMP-2/CRM *vs.* ICBG	38	36	53	53	40	12	FS, reoperation, OT, BL
Burkus et al, 2005[37]	RCT	Lumbar DDD	ALIF	rhBMP-2/ACS *vs.* ICBG	79	52	40.2	43.6	8.4–12	24	FS, ODI, SF-36, reoperation, NRS, OT, LOS, BL
Vaccaro et al, 2005[38]	RCT	Degenerative spondylolisthesis and stenosis	PLIF	rhBMP-7/Type 1 bone collagen *vs.* ICBG	24	12	63	66	7	24	FS, ODI, SF-36, reoperation,AD, OT, LOS, BL
Dimar et al, 2006[14]	RCT	Lumbar DDD, grade I spondylolisthesis	PLF	rhBMP-2/CRM *vs.* ICBG	53	45	50.9	52.7	40	24	FS, ODI, SF-36, reoperation, AD, NRS, OT, LOS, BL
Kanayama et al, 2006[33]	RCT	Degenerative spondylolisthesis with stenosis	PLIF	rhBMP-7/Type 1 bone collagen *vs.* ICBG	9	10	70.3	58.7	7	16.4	FS
Glassman et al, 2008[31]	RCT	Lumbar DDD or spondylolisthesis or stenosis;	PLF	rhBMP-2/ACS *vs.* ICBG	50	52	69.9	69.2	NR	24	FS, ODI, SF-36, reoperation, AD, NRS, OT, LOS, BL
Vaccaro et al, 2008[32]	RCT	Degenerative spondylolisthesis and stenosis	PLIF	rhBMP-7/Type 1 bone collagen *vs.* ICBG	257	87	68	69	7	>48	FS, ODI, SF-36, reoperation, AD, VAS, OT, LOS, BL
Dimar et al, 2009[29]	RCT	Lumbar DDD, grade I spondylolisthesis	PLF	rhBMP-2 matrix *vs.* ICBG	239	224	53.2	52.3	40	24	FS, ODI, SF-36, reoperation, AD, NRS, OT, LOS, BL
Dawson et al, 2009[30]	RCT	lumbar DDD, grade I spondylolisthesis	PLF	rhBMP-2/ACS *vs.* ICBG	25	21	55.9	56.9	12	24	FS, ODI\, SF-36, reoperation, AD, NRS, OT, LOS, BL
Delawi et al, 2010[28]	RCT	Grade II spondylolisthesis	PLIF	rhBMP-7/local autograft *vs.* ICBG	18	16	53	55	3.5	12	FS, ODI, AD, VAS, OT, LOS, BL
Michielsen et al, 2013[27]	RCT	Lumbar DDD	PLIF	rhBMP-2/ACS *vs.* ICBG	19	19	43.2	42.2	8	24	FS, ODI, SF-36, reoperation, AD, VAS, OT, LOS, BL
Hurlbert 2013[26]	RCT	Lumbar DDD	PLF	rhBMP-2/BCP *vs.* ICBG	98	99	53	53	42-63	48	FS, ODI, SF-36, reoperation, AD, NRS, OT, LOS, BL
Delawi et al, 2016[25]	RCT	Degenerative or isthmic spondylolisthesis ≤ grade 2	PLIF	rhBMP-7/local autograft *vs.* ICBG	60	59	54	55	7	24	FS, ODI, reoperation, AD, OT, LOS, BL

ACS indicates absorbable collagen sponge carrier; AD, adverse events; ALIF, anterior lumbar interbody fusion; BCP, biphasic calcium phosphate; BL, blood loss; BMP, bone morphogenetic protein; CRM, compression resistant matrix; DDD, degenerative disc disease; FS, fusion success; ICBG, autogenous iliac crest bone graft ; LOS, length of hospital stay; NR, no report; NRS, numeric rating scale for back and/or leg pain; ODI, Oswestry disability index; OT, operative time; PLF, posterior or posterolateral lumbar fusion; PLIF, posterior lumbar interbody fusion.

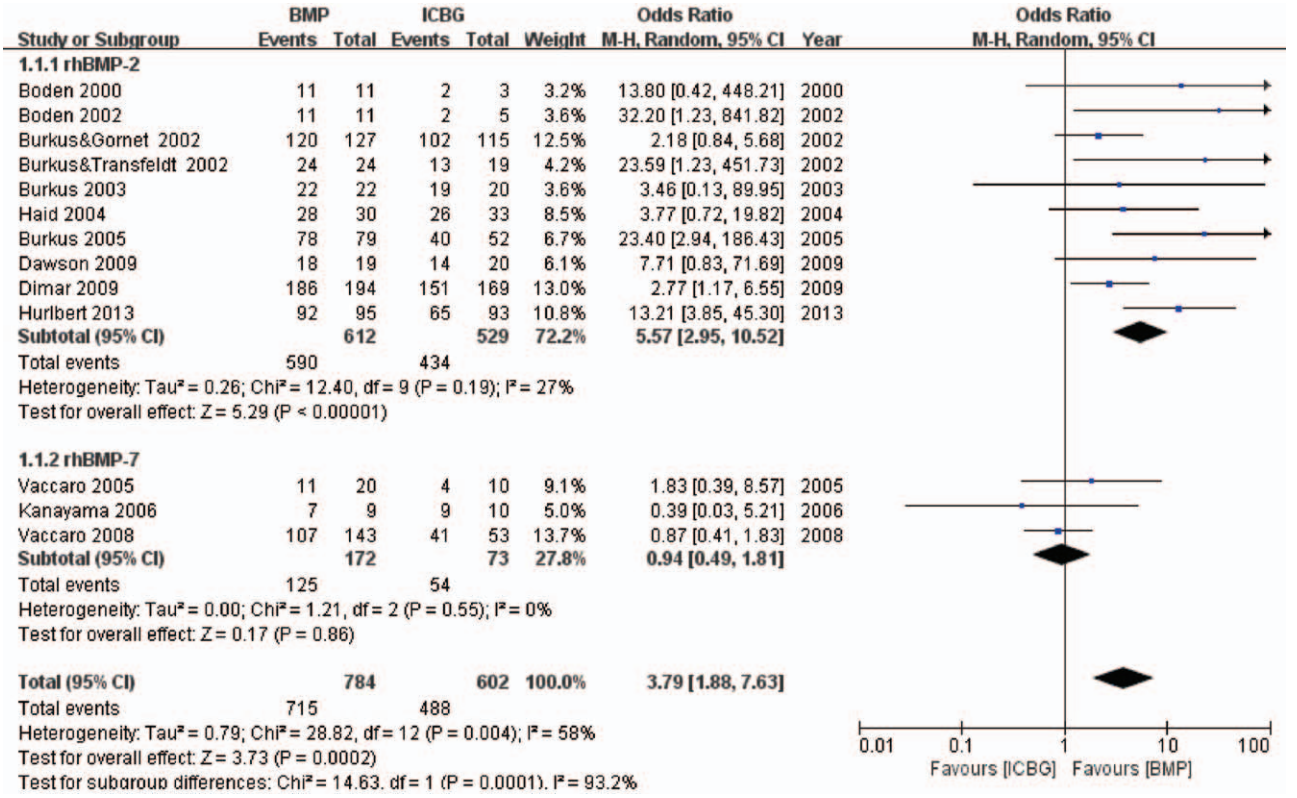

Figure 2. Fusion success—recombinant human bone morphogenetic protein versus autologous iliac bone graft.

Statistical analysis did not reveal significant differences between rhBMP and ICBG (OR 0.91, 95% CI 0.70–1.18, Figure 5). Different surgical approaches may affect the incidence of complications; therefore, subgroup analysis based on surgical procedure (ALIF or PLIF/PLF) was performed. No significant differences in ALIF (OR 0.78, 95% CI 0.37–1.64)[10,42,44] or PLIF/PLF (OR 0.93, 95% CI 0.71–1.22)[25–32,38,39,43] were observed when comparing rhBMP with ICBG. A funnel plot of the documented adverse events is presented in Figure 6. No evidence of publication bias was found. Besides, it is well-known that infection is an important complication following implant surgery. Therefore, we paid special attention to the data in this regard. Four studies reported surgical infections, involving 295

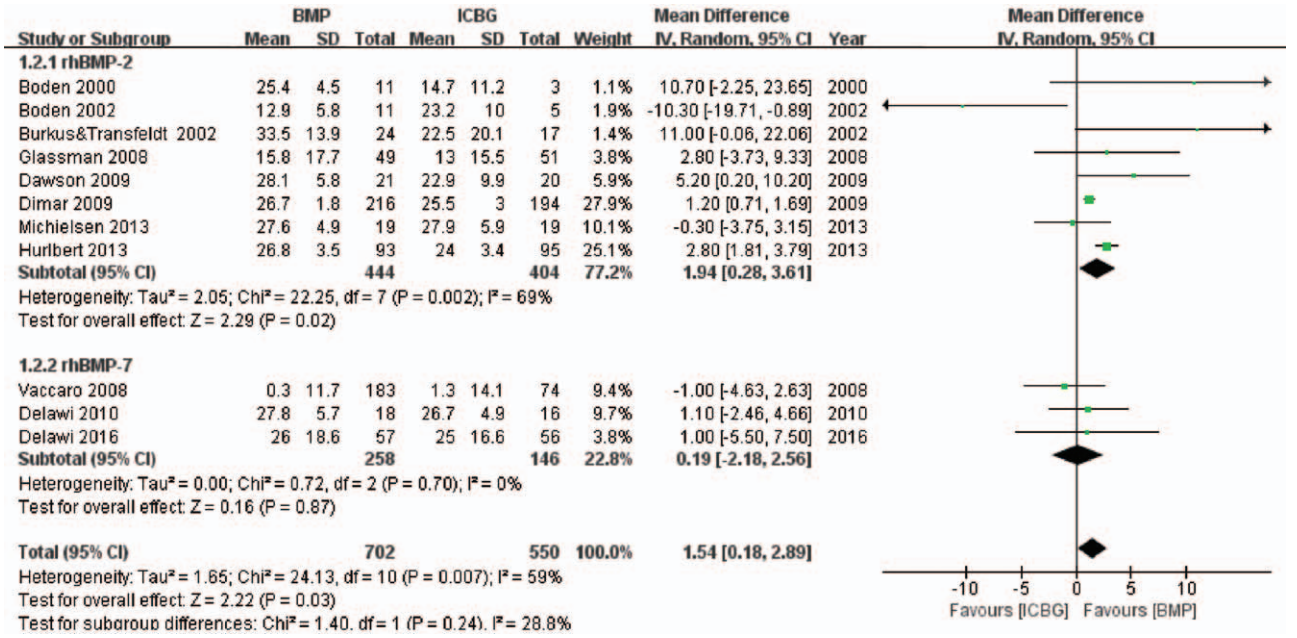

Figure 3. Improvement on ODI scores—recombinant human bone morphogenetic protein versus autologous iliac bone graft.

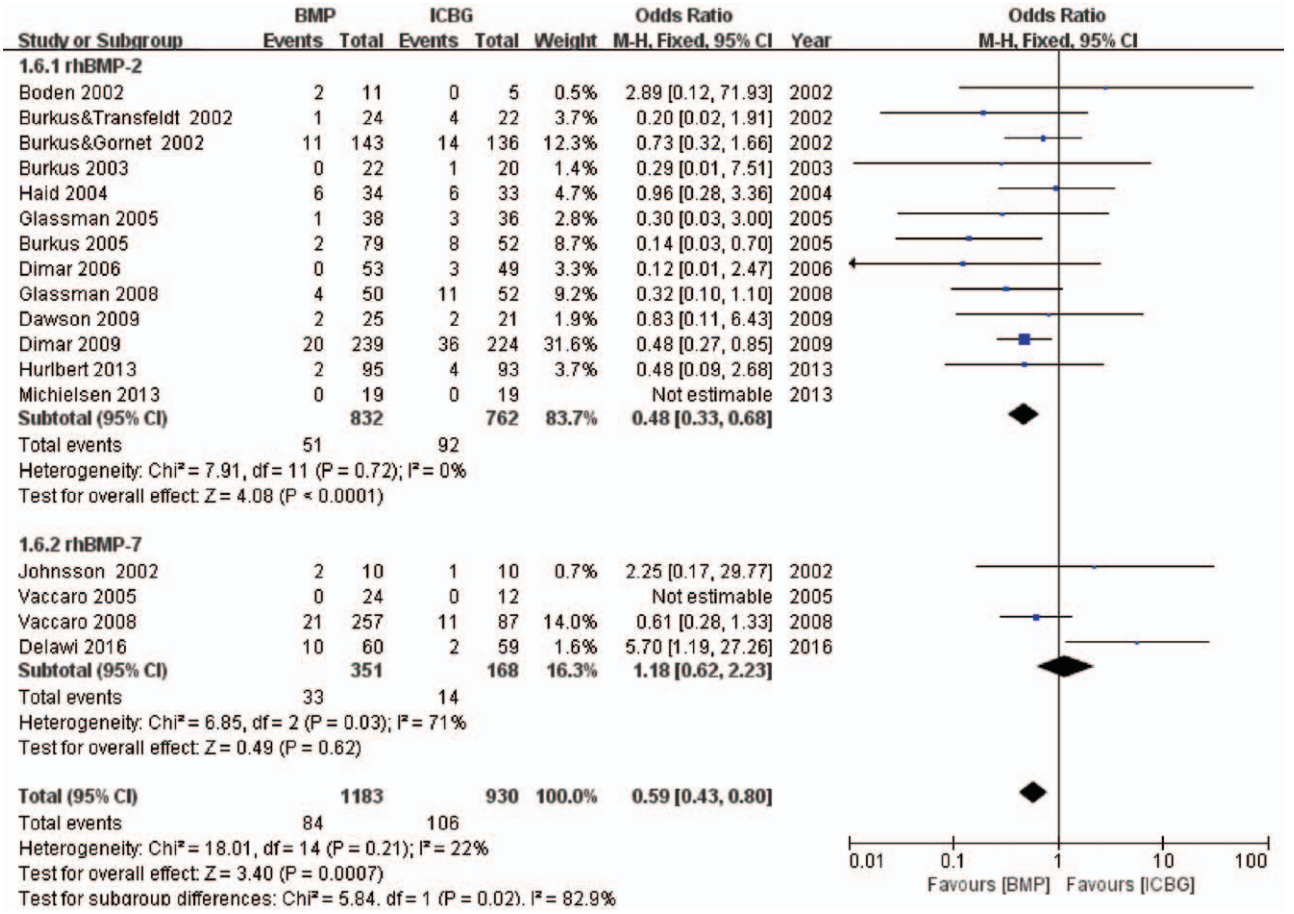

Figure 4. Reoperation—recombinant human bone morphogenetic protein versus autologous iliac bone graft.

patients. Two[30,31] of these studies were about rhBMP-2, and two[25,28] more were about rhBMP-7. No significant difference was identified between rhBMP and ICBG (OR 0.76, 95% CI 0.29–2.00). Subgroup analysis identified no significant differences in rhBMP-2 (OR 0.37, 95% CI −0.07 to 1.99) or rhBMP-7 (OR 1.17, 95% CI 0.34–4.02) were observed when compared with ICBG.

Secondary Outcomes

Operation Time

Concerning operation time, 10 studies were analyzed. Seven[10,26,27,29–31,42] of these studies included 876 patients and were about rhBMP-2, and three[25,28,38] more including 189 patients were about rhBMP-7. An overall reduced operation time was observed in rhBMP compared to ICBG (MD, −0.23, 95%CI, −0.44 to −0.02, Figure 7). Subgroup analysis revealed a significant difference between rhBMP-2 and ICBG (MD −0.27, 95% CI −0.54 to −0.01), with no significant difference observed between rhBMP-7 compared with ICBG (MD −0.13, 95% CI −0.38 to 0.11).

Intraoperative Blood Loss

No significant difference in blood loss was identified between rhBMP and ICBG in nine studies (MD −27.21, 95% CI −80.53 to 26.12). Seven[10,26,27,29–31,42] of these studies including 876 patients were about rhBMP-2, and two[25,28] more including 155 patients were about rhBMP-7. Subgroup analysis identified no significant differences in rhBMP-2 (MD −34.89, 95% CI −97.29 to 27.52) or rhBMP-7 (MD 6.32, 95% CI −84.41 to 97.05) when comparing with ICBG.

Duration of Hospital Stay

Nine studies provided specific data regarding hospital stay and reported no significant difference between rhBMP and ICBG (MD −0.52, 95% CI −1.02 to −0.01). Six[10,26,29–31,42] of these studies including 838 patients were about rhBMP-2, and two[25,28,38] more including 187 patients were about rhBMP-7. Subgroup analysis revealed a significant difference between rhBMP-2 and the ICBG (MD −0.64, 95% CI −1.22 to −0.06), with no significant difference observed between rhBMP-7 compared with ICBG (MD, 0.07, 95% CI −0.98 to 1.12).

Risk of Bias in Included Studies

Following assessment of all original studies, we evaluated bias risk using the 12 criteria recommended by the Cochrane Back Review Group.[22] The migration risk for each study is described in Table 2. The ratings across all included studies were summarized and presented in Figure 8.

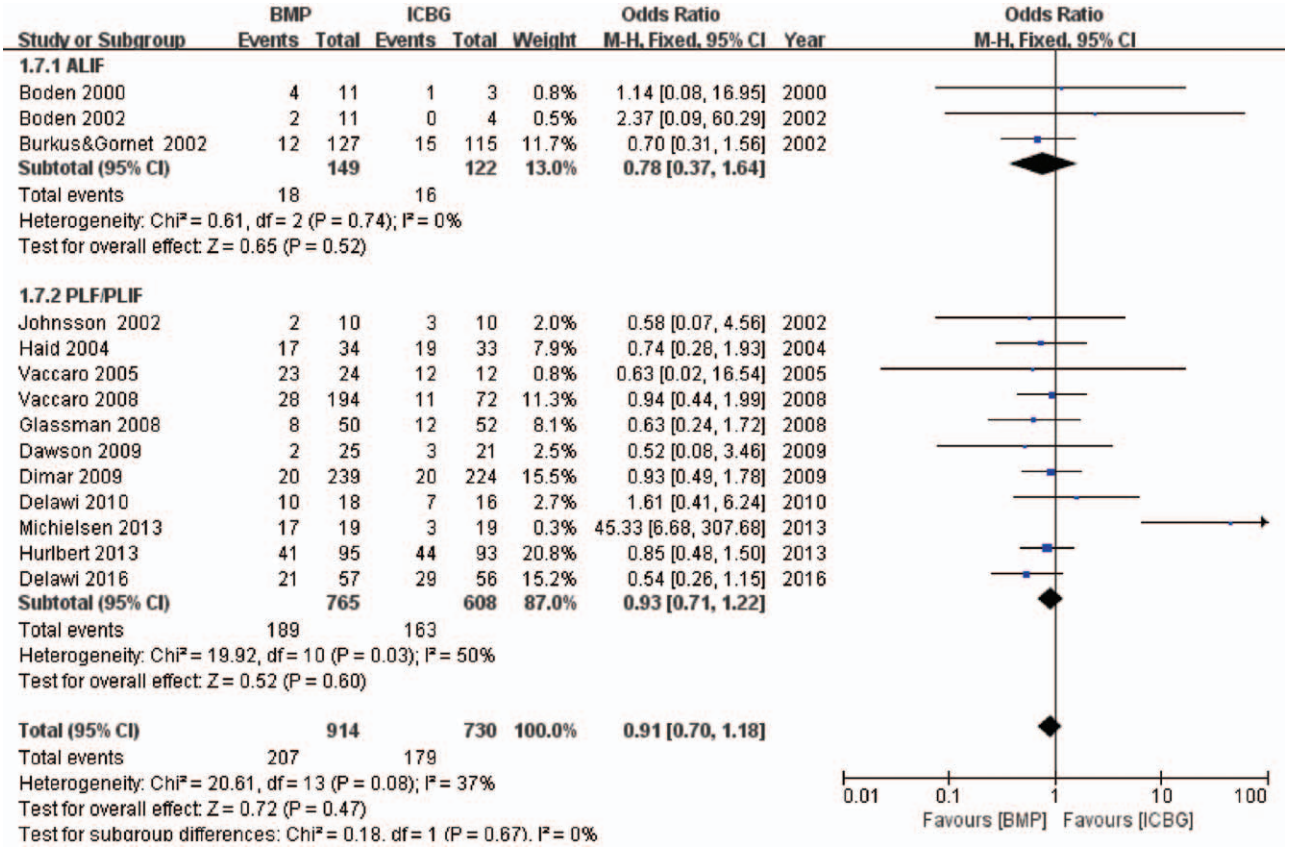

Figure 5. Adverse events—recombinant human bone morphogenetic protein versus autologous iliac bone graft.

Sensitivity Analysis and Quality Assessment

We performed a sensitivity analysis of the primary outcomes that exhibited higher heterogeneity. The primary method we utilized was to eliminate low-quality research on a study-by-study basis. When the studies of Hurlbert *et al*[26] were excluded, heterogeneity was significantly reduced. One explanation may be that the study's end point time was 4 years, with most other studies ending at 2 years. This indicates that the modifications in these indicators are related to the length of follow-up. GRADE quality assessment was used to rate the quality of evidence for all pooled outcomes. These results are shown in Table 2.

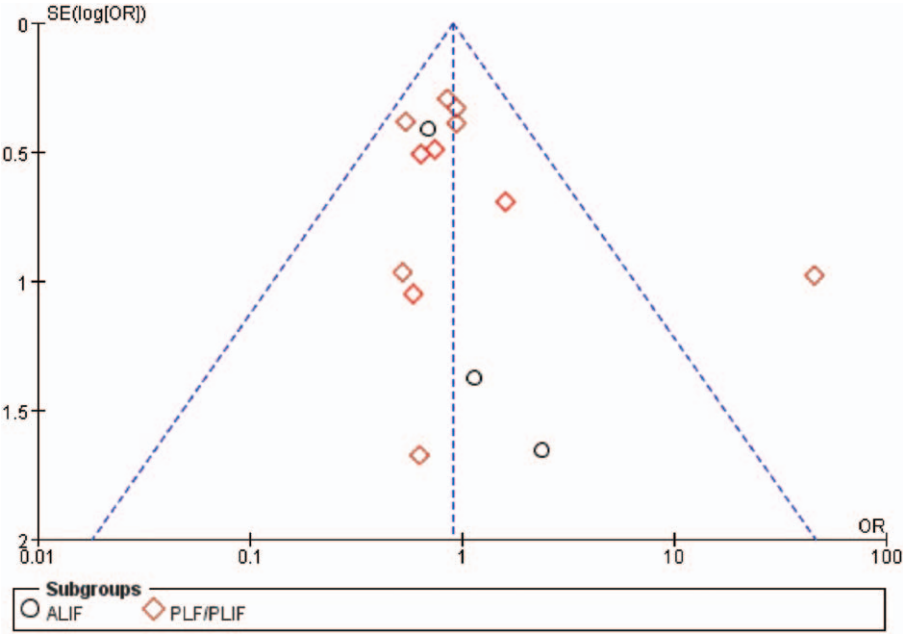

Figure 6. Funnel plot—adverse events.

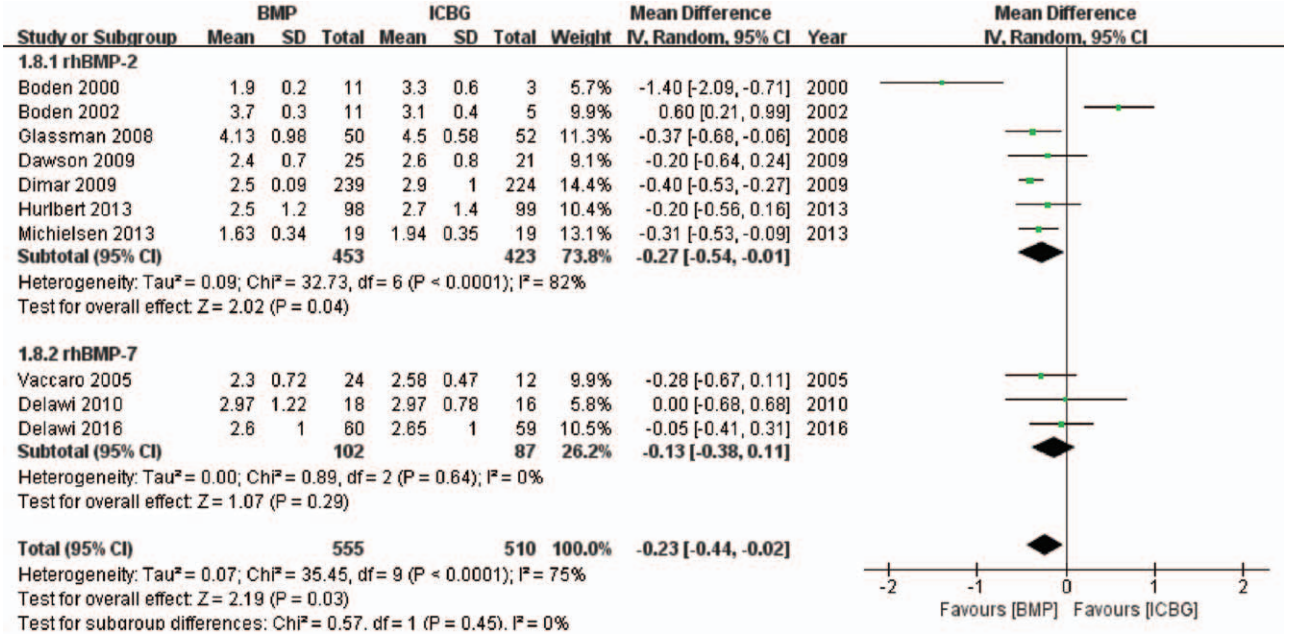

Figure 7. Operation time—recombinant human bone morphogenetic protein versus autologous iliac bone graft.

TABLE 2. Summary of Findings: Meta-analysis Comparison of Bone Morphogenic Protein and Iliac Crest in Degenerative Lumbar Disease

Outcomes	Anticipated Absolute Effects* (95% CI)		Relative effect (95% CI)	No. of Participants (Studies)	Certainty of the Evidence (Grade)	Conclusion
	Risk With Autologous Iliac Crest Bone Graft	Risk With Recombinant Human Bone Morphogenetic Protein				
Fusion success	811 per 1000	942 per 1000 (889–970)	OR 3.79 (1.88–7.63)	1386 (13 RCTs)	⊕⊕◯◯ Low†,‡	rhBMP may improve fusion rate
Fusion success—rhBMP-2	820 per 1000	962 per 1000 (931–980)	OR 5.57 (2.95–10.52)	1141 (10 RCTs)	⊕⊕⊕◯ moderate§	rhBMP-2 likely increases fusion rate
Fusion success—rhBMP-7	740 per 1000	728 per 1000 (582–837)	OR 0.94 (0.49–1.81)	245 (3 RCTs)	⊕⊕⊕◯ moderate‖	rhBMP-7 may not improve fusion rate
Improvement of ODI	The mean ODI ranged from 1.3 to 27.9¶	The mean ODI in the intervention group was 1.54 higher (0.18 higher to 2.89 higher)	—	1252 (11 RCTs)	⊕⊕◯◯ low‡,#	rhBMP may improve ODI
Improvement of SF-36 (PCS)	The mean SF-36 (PCS) ranged from 7 to 16.5¶	The mean SF-36 (PCS) in the intervention group was 1.16 higher (0.84 lower to 3.16 higher)	—	564 (4 RCTs)	⊕⊕◯◯ low‡,**	rhBMP does not impact SF-36 (PCS)
Improvement of NRS back pain	The mean NRS back pain ranged from 3.4 to 8.1¶	The mean NRS back pain in the intervention group was 0.33 higher (0.51 lower to 1.18 higher)	—	567 (4 RCTs)	⊕⊕◯◯ low‡,††	rhBMP does not impact NRS-back pain

TABLE 2 (*Continued*)

Outcomes	Anticipated Absolute Effects* (95% CI)		Relative effect (95% CI)	No. of Participants (Studies)	Certainty of the Evidence (Grade)	Conclusion
	Risk With Autologous Iliac Crest Bone Graft	Risk With Recombinant Human Bone Morphogenetic Protein				
Improvement of NRS leg pain	The mean NRS leg pain ranged from 3.1 to 7.3¶	The mean NRS leg pain in the intervention group was 0.84 higher (0.02 higher to 1.65 higher)	—	755 (5 RCTs)	⨁⨁◯◯ low‡,††	rhBMP may improve NRS leg pain
Re-operation	114 per 1000	71 per 1000 (52–93)	OR 0.59 (0.43–0.80)	2113 (17 RCTs)	⨁⨁⨁◯ moderate‡‡	rhBMP may reduce reoperation
Adverse events	245 per 1000	228 per 1000 (185–277)	OR 0.91 (0.70–1.18)	1644 (14 RCTs)	⨁⨁⨁◯ moderate§§	rhBMP may not reduce adverse events

Patient or population: conservative treatment of lumbar degenerative diseases invalid patients.

Setting: lumbar fusion.

Intervention: recombinant human bone morphogenetic protein.

Comparison: autologous iliac bone graft.

CI indicates confidence interval; OR, odds ratio; MD, mean difference; NRS, Numeric Rating Scale; ODI, Oswestry disability index; PCS, Physical Component Summary; RCT, randomized controlled trial; rhBMP, recombinant human bone morphogenetic protein; SF-36, short form-36.

GRADE Working Group grades of evidence.

High certainty: We are very confident that the true effect lies close to that of the estimate of the effect.

Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.

Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

**The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).*

Explanations:

†Eight studies did not describe the randomization process. twelve studies did not describe allocation hiding. All study was not blinded to the surgeon or patient.

‡High heterogeneity.

§Six studies did not describe the randomization process. Nine studies did not describe allocation hiding. All study was not blinded to the surgeon or patient.

||Two studies did not describe the randomization process. All studies did not describe allocation hiding. All study was not blinded to the surgeon or patient.

¶The mean ODI/SF-36(PCS)/NRS back pain/NRS leg pain is the means of improvement between preoperative and postoperative outcome.

#Six studies did not describe the randomization process. Eight studies did not describe allocation hiding. All study was not blinded to the surgeon or patient.

***All studies did not describe the randomization process. Three studies did not describe allocation hiding. All study was not blinded to the surgeon or patient.*

††All studies did not describe the randomization process. All studies did not describe allocation hiding. All study was not blinded to the surgeon or patient.

‡‡Eleven studies did not describe the randomization processThirteen studies did not describe allocation hiding. All study was not blinded to the surgeon or patient.

§§Nine studies did not describe the randomization process. Ten studies did not describe allocation hiding. All study was not blinded to the surgeon or patient.

DISCUSSION

Lumbar fusion with ICBG is the criterion standard surgical procedure for discogenic pain refractory to conservative treatments.[46–48] However, due to some well-known complications, some bone graft substitutes including rhBMP have been developed and applied in clinical practice. Five systematic reviews[16,49–52] suggesting that rhBMP reduced risk of fusion failure compared with ICBG, but had no advantage over ICBG concerning pain relief and functional recovery. Fu *et al* and Simmonds *et al* suggested that there was some evidence that rhBMP-2 may cause serious complications.[49,50] Zhang *et al*[15] included 19 RCTs for systematic review, with results concluding that rhBMP can reduce the reoperation rate and operation time. However, primary outcomes in that study such as fusion success rate and clinical success, were not clearly defined. Ye *et al*[17] suggested that using rhBMP-7 instead of ICBG produced no any additional benefits in single level PLF. Han *et al*[53] reported no significant differences in fusion rate of no internal fixation at 24 months for rhBMP, but there was high risk of bias due to including pooling nonrandomized controlled trials and lacking subgroup analysis. Recently Mariscal *et al*[54] found rhBMP-2 in PLF reduced surgical morbidity and had more beneficial effects on the fusion rate. However, multiple important

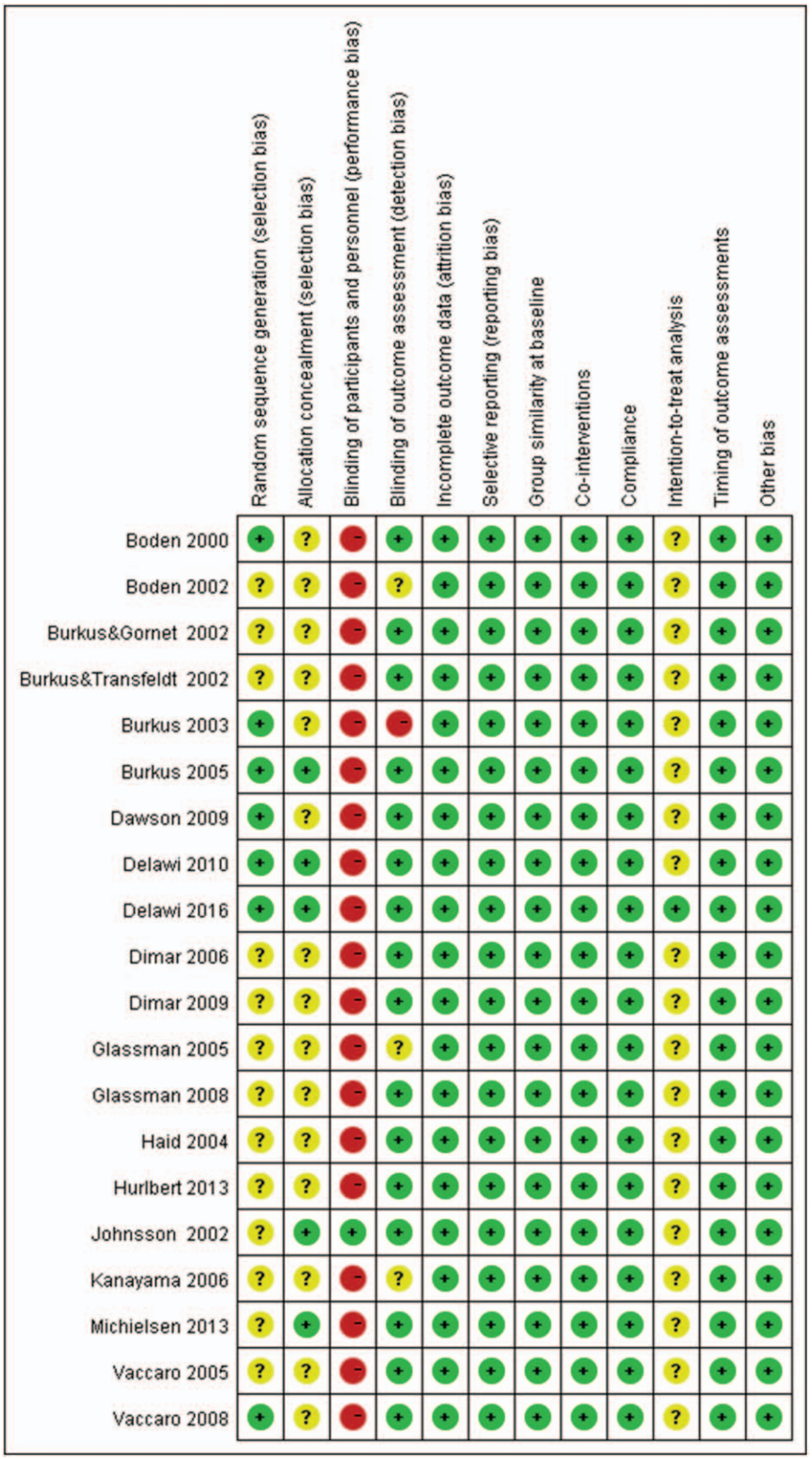

Figure 8. Risk of bias summary. This risk of bias tool incorporates assessment of randomization (sequence generation and allocation concealment), blinding (participants, personnel and outcome assessors), completeness of outcome data, selection of outcomes reported, and other sources of bias. The items were scored with ''yes,'' ''no,'' or ''unclear.''

RCTs[10,14,27,35,37,39,41,44,45] were not included in the analysis. The quality of evidence was not assessed for the included study and pooled results, and sensitivity analysis was not performed. As a result, evidence is still lacking to support the superiority of rhBMP compared with ICBG. Our present study serves as an update the in systematic evaluation. We incorporated the latest two large-sample RCTs,[25,26] utilizing a systematic review method approved by Cochrane

Collaboration. This included extracting data more carefully, performing sensitivity analyses for pooled outcomes, assessing risk of bias and quality of evidence according to the GRADE approach. Our attention was focused not only on subgroup analysis, which based on protein type. Data processing of complications was also a subgroup analysis emphasizing anterior and posterior surgery.

We determine that rhBMP is a fusion material that demonstrates better efficacy compared to ICBG. In terms of fusion success, rhBMP was approximately 2.8 times more effective than ICBG. Further analysis showed that the fusion success rate in rhBMP-2 was approximately 5.5 times higher than that observed in ICBG, whereas the fusion rate with rhBMP-7 was 5% less than ICBG. Similarly, the ratio of reoperation in rhBMP was about 60% of ICBG. Subgroup analysis showed that the rhBMP-2 had reduced incidence of reoperation by approximately 40% compared with ICBG, whereas rhBMP-7 had approximately 1.2 times the incidence than ICBG. These results indirectly suggest that rhBMP-2 is more effective than rhBMP-7 in inducing bone formation. This difference may be due to the different carriers used, and it is necessary to perform a RCT comparing rhBMP-2 and rhBMP-7 efficacy to confirm our conclusions. A summary of clinical treatment effects indicates a difference in ODI and NRS leg pain, and rhBMP-2 appears to be a better approach.

In terms of safety, whether using ALIF or PLIF/PLF, there was no significant difference in incidence of complications between rhBMP and ICBG. It is worth mentioning that surgical site infection should be taken seriously. Delawi *et al*[25] reported that four patients required reoperation due to surgical infection. This serves as a reminder that it is necessary to pay attention to the surgical incision in the early postoperative period. In addition, it is important to treat surgical site infection as early and aggressively as possible to prevent a more serious infection that requires antibiotic treatment.[30] Regarding surgical data, operation time and hospitalization days with rhBMP-2 treatment is reduced compared to ICBG. But, with respect to blood loss, rhBMP-2 does not show much advantage. What is different from previous study[17] is that there was no significant difference in any surgical data including hospital stay between rhBMP-7 and ICBG.

Despite the complexity of our study, we recognize that it still has some limitations. First, although there were 20 studies in total, a large portion of the studies did not provide SD values, and the inability to extract valid data led to limited data accuracy. In addition, the quality of evidence in this meta-analysis is limited by the low quality of the original studies. Most evaluated studies did not report their randomization or allocation methods. Nearly all studies failed to use independent blinding. Finally, a no cost–benefit analysis was conducted. Aside from the cost of research by Glassman *et al*,[31] other studies have not reported this parameter, which made analysis difficult or impossible. Therefore, subsequent RCT should pay attention to the application of more rigorous methods and indicators, including more accurate reporting of pre- and postoperative scores, follow-up of long-term complications, and costs of treatment.

Our review indicates that rhBMP-2 may be superior in terms of fusion success, ODI, reoperation, and duration of hospital stay. It might represent a suitable substitute for ICBG in lumbar fusion. Conversely, rhBMP-7 is not recommended for lumbar fusion. Further studies including cost-effective data analysis and RCT for the comparison of the efficacy of rhBMP-2 with rhBMP-7 is necessary to confirm the results observed in the present study.

➢ Key Points

- ❑ RhBMP-2 seems to be a more effective fusion material than ICBG.
- ❑ RhBMP-2 may be superior in terms of fusion success, improvement of ODI, re-operation, operation time, and length of hospital stay, but rhBMP-7 has no significant advantage compared with ICBG.
- ❑ No significant differences were observed between rhBMP-2 and ICBG regarding improvement of SF-36 and NRS back pain, adverse events, and blood loss.

References

1. Barrey CY, Le Huec JC. Chronic low back pain: relevance of a new classification based on the injury pattern. *Orthopa Traumatol Surg Res* 2019;105:339–46.
2. Deyo RA, Nachemson A, Mirza SK. Spinal-fusion surgery—the case for restraint. *N Engl J Med* 2004;350:722–6.
3. Mobbs RJ, Phan K, Malham G, et al. Lumbar interbody fusion: techniques, indications and comparison of interbody fusion options including PLIF, TLIF, MI-TLIF, OLIF/ATP, LLIF and ALIF. *J Spine Surg (Hong Kong)* 2015;1:2–18.
4. Boden SD. Overview of the biology of lumbar spine fusion and principles for selecting a bone graft substitute. *Spine (Phila Pa 1976)* 2002;27:S26–31.
5. Arrington ED, Smith WJ, Chambers HG, et al. Complications of iliac crest bone graft harvesting. *Clin Orthop Relat Res* 1996;300–9.
6. Vaccaro AR, Chiba K, Heller JG, et al. Bone grafting alternatives in spinal surgery. *Spine J* 2002;2:206–15.
7. Urist MR, Strates BS. Bone morphogenetic protein. *J Dental Res* 1971;50:1392–406.
8. Celeste AJ, Iannazzi JA, Taylor RC, et al. Identification of transforming growth factor beta family members present in bone-inductive protein purified from bovine bone. *Proc Natl Acad Sci U S A* 1990;87:9843–7.
9. Ong KL, Villarraga ML, Lau E, et al. Off-label use of bone morphogenetic proteins in the United States using administrative data. *Spine (Phila Pa 1976)* 2010;35:1794–800.
10. Boden SD, Zdeblick TA, Sandhu HS, et al. The use of rhBMP-2 in interbody fusion cages. Definitive evidence of osteoinduction in humans: a preliminary report. *Spine (Phila Pa 1976)* 2000; 25:376–81.
11. Villavicencio AT, Burneikiene S, Nelson EL, et al. Safety of transforaminal lumbar interbody fusion and intervertebral recombinant human bone morphogenetic protein-2. *J Neurosurg Spine* 2005; 3:436–43.

12. Williams BJ, Smith JS, Fu KM, et al. Does bone morphogenetic protein increase the incidence of perioperative complications in spinal fusion? A comparison of 55,862 cases of spinal fusion with and without bone morphogenetic protein. *Spine (Phila Pa 1976)* 2011;36:1685–91.
13. Cahill KS, Chi JH, Day A, et al. Prevalence, complications, and hospital charges associated with use of bone-morphogenetic proteins in spinal fusion procedures. *JAMA* 2009;302:58–66.
14. Dimar JR, Glassman SD, Burkus KJ, et al. 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 (Phila Pa 1976)* 2006;31:2534–9; discussion 40.
15. Zhang H, Wang F, Ding L, et al. A meta analysis of lumbar spinal fusion surgery using bone morphogenetic proteins and autologous iliac crest bone graft. *PLoS One* 2014;9:e97049.
16. Noshchenko A, Hoffecker L, Lindley EM, et al. Perioperative and long-term clinical outcomes for bone morphogenetic protein versus iliac crest bone graft for lumbar fusion in degenerative disk disease: systematic review with meta-analysis. *J Spinal Disord Tech* 2014;27:117–35.
17. Ye F, Zeng Z, Wang J, et al. Comparison of the use of rhBMP-7 versus iliac crest autograft in single-level lumbar fusion: a meta-analysis of randomized controlled trials. *J Bone Miner Metab* 2018;36:119–27.
18. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine* 2000;25:2940–52.
19. McHorney CA, Ware JE, Lu JF, et al. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Medical care* 1994;32:40–66.
20. Farrar JT, Young JP, LaMoreaux L, et al. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* 2001;94:149–58.
21. Burkus JK, Foley K, Haid RW, et al. Surgical Interbody Research Group–radiographic assessment of interbody fusion devices: fusion criteria for anterior lumbar interbody surgery. *Neurosurgical focus* 2001;10:E11.
22. Furlan AD, Malmivaara A, Chou R, et al. 2015 Updated Method Guideline for Systematic Reviews in the Cochrane Back and Neck Group. *Spine (Phila Pa 1976)* 2015;40:1660–73.
23. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. *BMJ (Clinical research ed)* 2004;328:1490.
24. Cho JH, Lee JH, Yeom JS, et al. Efficacy of Escherichia coli-derived recombinant human bone morphogenetic protein-2 in posterolateral lumbar fusion: an open, active-controlled, randomized, multicenter trial. *Spine J* 2017;17:1866–74.
25. Delawi D, Jacobs W, van Susante JL, et al. OP-1 compared with iliac crest autograft in instrumented posterolateral fusion: a randomized, multicenter non-inferiority trial. *J Bone Joint Surg Am* 2016;98:441–8.
26. Hurlbert RJ, Alexander D, Bailey S, et al. rhBMP-2 for posterolateral instrumented lumbar fusion: a multicenter prospective randomized controlled trial. *Spine (Phila Pa 1976)* 2013;38:2139–48.
27. Michielsen J, Sys J, Rigaux A, et al. The effect of recombinant human bone morphogenetic protein-2 in single-level posterior lumbar interbody arthrodesis. *J Bone Joint Surg Am* 2013;95:873–80.
28. Delawi D, Dhert WJ, Rillardon L, et al. A prospective, randomized, controlled, multicenter study of osteogenic protein-1 in instrumented posterolateral fusions: report on safety and feasibility. *Spine (Phila Pa 1976)* 2010;35:1185–91.
29. Dimar JR, Glassman SD, Burkus JK, et al. Clinical and radiographic analysis of an optimized rhBMP-2 formulation as an autograft replacement in posterolateral lumbar spine arthrodesis. *J Bone Joint Surg Am* 2009;91:1377–86.
30. Dawson E, Bae HW, Burkus JK, et al. Recombinant human bone morphogenetic protein-2 on an absorbable collagen sponge with an osteoconductive bulking agent in posterolateral arthrodesis with instrumentation. A prospective randomized trial. *J Bone Joint Surg Am* 2009;91:1604–13.
31. Glassman SD, Carreon LY, Djurasovic M, et al. RhBMP-2 versus iliac crest bone graft for lumbar spine fusion: a randomized, controlled trial in patients over sixty years of age. *Spine (Phila Pa 1976)* 2008;33:2843–9.
32. Vaccaro AR, Lawrence JP, Patel T, et al. The safety and efficacy of OP-1 (rhBMP-7) as a replacement for iliac crest autograft in posterolateral lumbar arthrodesis: a long-term (>4 years) pivotal study. *Spine (Phila Pa 1976)* 2008;33:2850–62.
33. Kanayama M, Hashimoto T, Shigenobu K, et al. 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 (Phila Pa 1976)* 2006;31:1067–74.
34. Burkus JK, Sandhu HS, Gornet MF. Influence of rhBMP-2 on the healing patterns associated with allograft interbody constructs in comparison with autograft. *Spine (Phila Pa 1976)* 2006;31:775–81.
35. Glassman SD, Dimar JR, Carreon LY, et al. 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 (Phila Pa 1976)* 2005;30:1694–8.
36. 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.
37. Burkus JK, Sandhu HS, Gornet MF, et al. Use of rhBMP-2 in combination with structural cortical allografts: clinical and radiographic outcomes in anterior lumbar spinal surgery. *J Bone Joint Surg Am* 2005;87:1205–12.
38. Vaccaro AR, Anderson DG, Patel T, et al. Comparison of OP-1 Putty (rhBMP-7) to iliac crest autograft for posterolateral lumbar arthrodesis: a minimum 2-year follow-up pilot study. *Spine (Phila Pa 1976)* 2005;30:2709–16.
39. Haid RW, Branch CL, Alexander JT, et al. Posterior lumbar interbody fusion using recombinant human bone morphogenetic protein type 2 with cylindrical interbody cages. *Spine J* 2004;4:527–38; discussion 38-9.
40. Vaccaro AR, Patel T, Fischgrund J, et al. A pilot study evaluating the safety and efficacy of OP-1 Putty (rhBMP-7) as a replacement for iliac crest autograft in posterolateral lumbar arthrodesis for degenerative spondylolisthesis. *Spine (Phila Pa 1976)* 2004;29:1885–92.
41. Burkus JK, Dorchak JD, Sanders DL. Radiographic assessment of interbody fusion using recombinant human bone morphogenetic protein type 2. *Spine (Phila Pa 1976)* 2003;28:372–7.
42. Boden SD, Kang J, Sandhu H, et al. 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)* 2002;27:2662–73.
43. Johnsson R, Strömqvist B, Aspenberg P. Randomized radiostereometric study comparing osteogenic protein-1 (BMP-7) and autograft bone in human noninstrumented posterolateral lumbar fusion: 2002 Volvo Award in clinical studies. *Spine (Phila Pa 1976)* 2002;27:2654–61.
44. Burkus JK, Gornet MF, Dickman CA, et al. Anterior lumbar interbody fusion using rhBMP-2 with tapered interbody cages. *J Spinal Disord Tech* 2002;15:337–49.
45. Burkus JK, Transfeldt EE, Kitchel SH, et al. Clinical and radiographic outcomes of anterior lumbar interbody fusion using recombinant human bone morphogenetic protein-2. *Spine* 2002;27:2396–408.
46. Kaufman HH, Jones E. The principles of bony spinal fusion. *Neurosurgery* 1989;24:264–70.
47. Fujiya M, Saita M, Kaneda K, et al. Clinical study on stability of combined distraction and compression rod instrumentation with posterolateral fusion for unstable degenerative spondylolisthesis. *Spine (Phila Pa 1976)* 1990;15:1216–22.

48. Blumenthal SL, Baker J, Dossett A, et al. The role of anterior lumbar fusion for internal disc disruption. *Spine (Phila Pa 1976)* 1988;13:566–9.
49. Fu R, Selph S, McDonagh M, et al. Effectiveness and harms of recombinant human bone morphogenetic protein-2 in spine fusion: a systematic review and meta-analysis. *Ann Intern Med* 2013;158:890–902.
50. Simmonds MC, Brown JV, Heirs MK, et al. Safety and effectiveness of recombinant human bone morphogenetic protein-2 for spinal fusion: a meta-analysis of individual-participant data. *Ann Intern Med* 2013;158:877–89.
51. Chen Z, Ba G, Shen T, et al. Recombinant human bone morphogenetic protein-2 versus autogenous iliac crest bone graft for lumbar fusion: a meta-analysis of ten randomized controlled trials. *Arch Orthop Trauma Surg* 2012;132:1725–40.
52. Papakostidis C, Kontakis G, Bhandari M, et al. 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:E680–92.
53. Han PF, Chen TY, Zhang ZL, et al. rhBMP in lumber fusion for lumbar spondylolisthesis: a systematic review and meta-analysis. *Chin J Traumatol* 2019;22:51–8.
54. Mariscal G, Nunez JH, Barrios C, et al. A meta-analysis of bone morphogenetic protein-2 versus iliac crest bone graft for the posterolateral fusion of the lumbar spine. *J BoneMiner Metab* 2019;22:51–8.