ORIGINAL ARTICLE

Efficacy and safety of recombinant human bone morphogenetic protein-2 biomaterials in promoting bone regeneration: A systematic review and meta-analysis

Yue Li, MD[1,2] ID, Jun Zhao, MD[1,2] ID, Chunxia Chen, MD[1,2] ID

[1]Department of Prosthodontics, Tianjin Stomatological Hospital, School of Medicine, Nankai University, Tianjin City, China
[2]Tianjin Key Laboratory of Oral and Maxillofacial Function Reconstruction, Tianjin City, China

Bone defect repair has long been a major problem in orthopedics and oral and maxillofacial surgery.[1,2] Although autologous bone transplantation is regarded as the gold standard, it is limited by complications such as limited bone supply, pain in the bone harvesting area and infection.[3-5] In recent years, bone morphogenetic protein-2 (BMP-2) has attracted clinical attention due to its potent osteogenic induction activity.[6] Animal experiments have shown that recombinant human (rh)BMP-2 can accelerate osteoblast differentiation and promote new bone formation.[7,8] Govender et al.[9] first demonstrated in an open tibial fracture randomized-controlled trial (RCT) that rhBMP-2 could shorten healing time. Subsequently, multiple RCTs have reported positive therapeutic effects in the fields of spinal

Citation: Li Y, Zhao J, Chen C. Efficacy and safety of recombinant human bone morphogenetic protein-2 biomaterials in promoting bone regeneration: A systematic review and meta-analysis. Jt Dis Relat Surg 2026;37(2):325-343. doi: 10.52312/jdrs.2026.2523.

ABSTRACT

Objectives: This study aims to systematically evaluate the efficacy and safety of recombinant human bone morphogenetic protein-2 (rhBMP-2) in promoting bone regeneration.

Materials and methods: A comprehensive search of PubMed, Embase, Web of Science, and Scopus was conducted from inception to May 2024. Fifteen randomized-controlled trials involving 2,137 cases were included. Traditional and network meta-analyses were performed, machine learning techniques were applied to explore heterogeneity, and subgroup analyses were carried out to assess efficacy across anatomical sites including alveolar sockets, palatal clefts, and spinal fusion.

Results: No significant difference was found in the overall bone healing rate between the rhBMP-2 and control groups (relative risk [RR] = 1.02, 95% confidence interval [CI] 0.87-1.20). However, rhBMP-2 demonstrated a significant advantage in spinal fusion (RR = 1.09, 95% CI 1.01-1.17) and a positive, although not statistically significant, trend in small oral and jaw bone defects. The incidence of serious adverse events was comparable (RR = 0.97, 95% CI 0.66-1.42). Network meta-analysis indicated that the overall success rate of bone regeneration with rhBMP-2 (odds ratio [OR] = 1.53, 95% CI 0.54-4.33) and other bone substitutes (OR = 1.42, 95% CI 0.39-5.21) did not significantly exceed autograft treatment, although both showed a trend toward superiority. Direct pairwise comparison revealed that rhBMP-2 was significantly more effective than autograft (OR = 1.54, 95% CI 1.02-2.33, $p < 0.05$), with low heterogeneity ($I^2 = 0\%$).

Conclusion: While the overall healing rate was comparable to controls, rhBMP-2 showed significant efficacy in spinal fusion and a favorable safety profile. Its effectiveness was context-specific, influenced by anatomical site and patient factors. These findings support the selective use of rhBMP-2 in specific clinical contexts such as spinal surgery, while further research is needed to optimize its application in other indications.

Keywords: Bone regeneration; heterogeneity, machine learning, network meta-analysis, recombinant human bone morphogenetic protein-2, safety, systematic review.

fusion,[10-12] alveolar ridge preservation,[13] and cleft palate repair.[14]

However, controversy over efficacy and safety has also emerged, with studies noting that high-dose rhBMP-2 may increase the risk of soft tissue swelling, heterotopic ossification and other complications.[14] The effect of different carrier materials on release kinetics also leads to significant differences in efficacy.[15,16] While previous systematic reviews have provided valuable insights within specific surgical indications, a broader synthesis is lacking. Many have not comprehensively accounted for the impact of critical variables such as carrier system and dosing regimen within their analyses. Furthermore, the existing body of literature lacks a unified analysis that quantifies the relative benefits of rhBMP-2 against a range of alternative treatment strategies across different clinical applications, leading to a fragmented evidence base.[17-19] Furthermore, a primary challenge in synthesizing the existing evidence is the high degree of heterogeneity, likely arising from the complex interplay of multiple factors, such as patient age, anatomical site, carrier material and dosage. Traditional methods for exploring heterogeneity, such as subgroup analysis and meta-regression, are limited in their ability to model complex, non-linear interactions between these covariates. This limitation often leads to an incomplete understanding of the sources of variability in treatment effects, contributing to the inconsistent conclusions across previous reviews.

To address this fundamental limitation, we incorporated machine learning (ML) algorithms into our meta-analytic framework. Unlike conventional statistical methods, ML models (e.g., random forest and gradient boosting) are specifically designed to handle high-dimensional, interacting and non-linear relationships without requiring pre-specified hypotheses. Therefore, the prominent advantage of ML in this context is its ability to objectively identify the most influential sources of heterogeneity from a wide array of candidate variables in a data-driven manner, even when their effects are complex and interdependent. This approach moves beyond testing pre-defined hypotheses to exploratory pattern recognition, offering a more robust and comprehensive method to explain why rhBMP-2 efficacy varies across studies. This constitutes a novel contribution to the existing literature on rhBMP-2 treatment.

Based on this, this study, under the premise of strictly following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) and Cochrane guidelines, combines network meta-analysis with ML for the first time to systematically compare the clinical efficacy of rhBMP-2, autologous bone and other synthetic substitute materials, and to explore the effects of dose, carrier type and patient characteristics on heterogeneity. Subgroup analyses are specifically conducted to evaluate the efficacy differences of rhBMP-2 treatment in different anatomical sites, such as oral and maxillofacial bone, spinal fusion and long bone fractures, aiming to provide an evidence-based foundation for clinical, individualized medication and future research design.

MATERIALS AND METHODS

Study design and registration

This study followed the standard methodological procedures for systematic reviews and network meta-analyses, and was designed and compiled in accordance with The Cochrane Handbook and the PRISMA 2020 statement.[20] The study protocol was pre-registered in advance to ensure the transparency of the methodology and the reproducibility of the results (INPLASY2025100054). As this study is a systematic review and meta-analysis using only previously published and de-identified aggregate data, it was exempt from institutional review board approval (Exemption Statement, Tianjin Stomatological Hospital, 25.12.2025).

Literature search strategy

The PubMed, Embase, Web of Science and Scopus databases were systematically searched between the inception of each database and May 2024. The search keywords included: 'recombinant human bone morphogenetic protein-2', 'rhBMP-2', 'BMP-2', 'bone regeneration', 'bone healing', 'fracture', 'spinal fusion', 'alveolar bone' and 'randomized controlled trial', and a comprehensive search strategy was formulated by combining subject terms and free terms. The search strategy was as follows:

PubMed (690): ('Bone Morphogenetic Protein 2'[Mesh]) OR ('BMP-2'[Title/Abstract]) OR ('rhBMP-2') OR ('bone morphogenetic protein-2') OR ('bone morphogenic protein-2') OR ('Recombinant human bone morphogenetic protein-2') AND ('Bone Regeneration'[Mesh] OR 'Bone regeneration' OR 'Bone healing' OR 'Fracture healing' OR 'Spinal fusion' OR 'Alveolar bone regeneration' OR 'Bone defect' OR 'Bone augmentation') AND (randomized

controlled trial) OR (controlled clinical trial[pt]) OR (randomized) OR (randomly) OR (trial).

Scopus (1,062): TITLE-ABS-KEY ('Bone Morphogenetic Protein 2' OR 'BMP-2' OR 'rhBMP-2' OR 'recombinant human bone morphogenetic protein-2') AND TITLE-ABS-KEY ('Bone regeneration' OR 'Bone healing' OR 'Fracture healing' OR 'Spinal fusion' OR 'Alveolar bone regeneration' OR 'Bone defect' OR 'Bone augmentation') AND TITLE-ABS-KEY ('randomized controlled trial' OR 'controlled clinical trial' OR 'randomized' OR 'randomly' OR 'trial').

Web of Science (1,080TS = ('Bone Morphogenetic Protein 2' OR 'BMP-2' OR 'rhBMP-2' OR 'recombinant human bone morphogenetic protein-2') AND TS = ('Bone regeneration' OR 'Bone healing' OR 'Fracture healing' OR 'Spinal fusion' OR 'Alveolar bone regeneration' OR 'Bone defect' OR 'Bone augmentation') AND TS = ('randomized' OR 'randomly' OR 'trial' OR 'controlled trial').

Embase (793): ('bone morphogenetic protein 2'/exp OR 'bone morphogenetic protein 2':ti,ab OR 'BMP-2':ti,ab' OR 'rhBMP-2':ti,ab' OR 'recombinant human bone morphogenetic protein-2':ti,ab') AND ('bone regeneration'/exp OR 'bone regeneration':ti,ab OR 'bone healing':ti,ab OR 'fracture healing':ti,ab OR 'spinal fusion':ti,ab OR 'alveolar bone regeneration':ti,ab OR 'bone defect':ti,ab OR 'bone augmentation':ti,ab) AND ('randomized controlled trial'/de OR 'randomized controlled trial':ti,ab OR 'controlled clinical trial':ti,ab OR 'randomized':ti,ab OR 'randomly':ti,ab OR 'trial':ti,ab).[21]

Inclusion and exclusion criteria

Inclusion criteria were as follows: (1) study type is RCT, no language restrictions; and (2) study population is human patients requiring bone regeneration (e.g., for critical-sized defects, spinal fusion alveolar ridge preservation) or fracture healing (e.g., traumatic long bone fractures).

Exclusion criteria were as follows: (1) non-randomized-controlled studies (e.g., case series, observational studies); (2) animal studies or basic experimental studies; (3) studies without complete efficacy outcome data; and (4) studies in duplicate publication or abstract form.

Intervention and outcome

Regarding the intervention measures, the experimental group received rhBMP-2 treatment (dosage form, dose and carrier type not limited).

The control measures included autologous bone transplantation or other bone replacement materials. The primary outcome indicators were imaging-based assessment of successful bone formation/union, measured as a bone regeneration success rate, fracture healing rate or spinal fusion rate. The secondary outcome indicator was serious adverse event (SAE) rate.

Literature screening and data extraction

Two researchers independently screened the literature, first based on the title and abstract. The full text of the studies that passed the initial screening was, then, screened. When disagreements occurred during the literature screening process, they were resolved through discussion involving a third-party researcher.

A standard data extraction form was used to extract the following characteristic information of each included study: basic information of the study (author, year, design type, follow-up duration); patient baseline information (age, smoking, diabetes and other comorbidities); intervention details (BMP-2 dose, carrier type, control type); and outcome indicator data (number of treatment events, number of adverse events, etc.). Following data extraction, cross-checking was performed to ensure accuracy and consistency.

Quality assessment and bias risk assessment

The Cochrane Collaboration's risk of bias tool (RoB 2.0, Cochrane Collaboration, Oxford, UK) was used to assess the risk of random sequence generation, allocation concealment, blinding of outcome assessment, incomplete outcome data and selective reporting bias.[22] The assessment was completed by two independent assessors, and consensus was reached through discussion or third-party arbitration when the assessment was inconsistent.

Statistical analysis

Meta-analysis was performed using the R software version 4.4.1 (R Foundation for Statistical Computing, Vienna, Austria) and the meta package for meta-analysis. For binary outcome indicators (e.g., bone regeneration success rate), the relative risk (RR) and its 95% confidence interval (CI) were used for combined effect size analysis. The heterogeneity evaluation used the I^2 test, with $I^2 > 50\%$ considered to indicate significant heterogeneity. When the heterogeneity was significant, the random-effects model was used; otherwise, the fixed-effects model was adopted. Funnel plots were created to assess the risk of publication bias.

Network meta-analysis was performed using the 'netmeta' package of the R software, and different controls (autologous bone, other bone substitute materials) were included in the same analysis framework. A network evidence map was created and the efficacy was ranked (surface under the cumulative ranking curve [SUCRA] value). The random-effects model was used for the analysis, and the odds ratio (OR) and 95% CI of each node treatment method relative to autologous bone were calculated to clarify the relative efficacy of different treatment strategies.

To explore the source of effect heterogeneity in meta-analysis, two ML algorithms, (random forest and gradient boosting) were further used to build models and analyze feature importance using Python (v3.10) [Python Software Foundation] and the scikit-learn library (v1.4.2) [scikit-learn development team]. Features included 14 items, such as total sample size, age, dose, carrier and study site. Based on the model prediction results, the characteristic factors that contributed most to heterogeneity were determined, and further visualization analysis was performed through the subgroup distribution of effect size.

Subgroup analysis was performed based on anatomical site (long bones, spine, maxillofacial region) and carrier type (absorbable collagen sponge [ACS], hyaluronic acid [HA], other synthetic materials) to test the stability of the effects of different clinical scenarios and materials on the treatment effect. In terms of safety, a meta-analysis was performed for SAEs reported in the study, and the combined RR and 95% CI were calculated to evaluate the safety of rhBMP-2 intervention. Publication bias was assessed visually using funnel plots, and the possibility of publication bias was quantitatively assessed using Egger's regression test. A p value of < 0.05 was considered statistically significant.

RESULTS

Literature screening process

The results of the study screening based on the PRISMA flowchart are shown in Figure 1. A total of 3,625 studies were identified following the database search. After removing 855 duplicate entries, 2,770 independent articles remained and were screened based on the title and abstract;
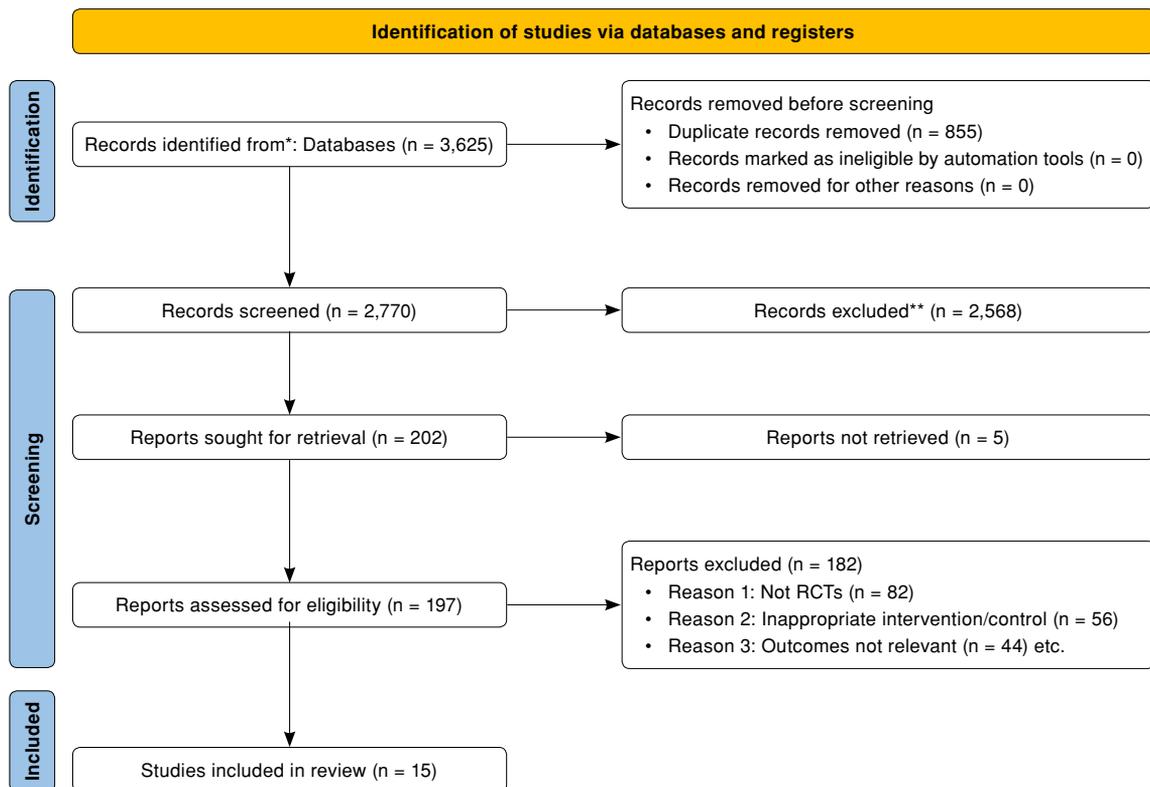
FIGURE 1. PRISMA literature screening flow chart. It demonstrates the process of literature screening, from database retrieval to the final inclusion of studies. Abbreviation:
PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; RCT, randomized-controlled trial; RR, risk ratio.

TABLE I

Characteristics of included literature

Study ID	Year	Treat node	Control node	Site	Design	Follow-up	N BMP	N CTRL	Dose total (mg)	Carrier
Govender et al.[9]	2002	BMP-2	Other	Tibia-open	Parallel	12	290	147	12	ACS
Aro et al.[23]	2011	BMP-2	Other	Tibia-open	Parallel	5	139	138	12	ACS
Lyon et al.[24]	2013	BMP-2	Other	Tibia-closed	Parallel	6	247	122	7.5	CPM
Haid et al.[27]	2004	BMP-2	Autograft	Lumbar PLIF	Parallel	24	34	33	6	ACS
Burkus et al.[25]	2003	BMP-2	Autograft	Lumbar ALIF	Parallel	24	143	136	6	ACS
Burkus et al.[26]	2002	BMP-2	Autograft	Lumbar ALIF	Parallel	24	143	136	6	ACS
Cho et al.[28]	2017	BMP-2	Autograft	Lumbar PLF	Parallel	6	42	51	3	HA
Choi et al.[29]	2022	BMP-2	Other	Lumbar PLF	Split-body	6	32	32	12.8	HA
Fiorellini et al.[30]	2005	BMP-2	Other	Alveolar socket	Parallel	4	40	40	1.9	ACS
Coomes et al.[31]	2014	BMP-2	Other	Alveolar socket	Parallel	5	20	19	2.1	ACS
Dickinson et al.[34]	2008	BMP-2	Autograft	Alveolar cleft	Parallel	12	9	12	6	ACS
Wei et al.[32]	2024	BMP-2	Other	Alveolar socket	Parallel	1.5	15	25	0.05	BioCaP-βTCP
Canan et al.[35]	2012	BMP-2	Autograft	Alveolar cleft	3-arm Parallel	12	6	6	3.6	ACS
Alonso et al.[36]	2010	BMP-2	Autograft	Alveolar cleft	Parallel	12	8	8	3.5	ACS
Jo et al.[33]	2019	BMP-2	Other	Alveolar socket	Parallel	4	32	32	1.05	ACS

BMP, bone morphogenetic protein; N BMP, number of patients in the BMP-2 treatment group; N CTRL, number of patients in the control group; ACS, absorbable collagen sponge; CPM, calcium phosphate matrix; PLIF, posterior lumbar interbody fusion; ALIF, anterior lumbar interbody fusion; PLF, posterolateral fusion; HA, hydroxyapatite; BioCaP, biomimetic calcium phosphate; β-TCP, beta-tricalcium phosphate.

2,568 articles were excluded because they did not meet the initial eligibility criteria related to study design, intervention or relevant outcomes. The full text of 202 studies was marked out for retrieval; however, five articles could not be retrieved, leaving 197 full-text articles for detailed eligibility evaluation. Following rigorous evaluation, 182 studies were excluded for the following reasons: 82 were not RCTs; 56 had inappropriate intervention or control conditions; and 44 did not report relevant outcomes. Finally, 15 studies were included in the qualitative synthesis and subsequent analysis, meeting all prespecified inclusion criteria and providing data on the efficacy and safety of rhBMP-2 biomaterials in bone regeneration.

Characteristics of included literature

This study included 15 RCTs published between 2002 and 2024, involving a total of 2,137 patients, including 1,200 patients who received rhBMP-2 intervention and 937 patients in the control group. The included studies involved three major clinical indications: tibial fractures (open and closed), spinal fusions (anterior and posterior) and maxillofacial bone defects. Studies on tibial fractures included two on open fractures (Govender et al.,[9]; Aro et al.,[23]) and one on closed fracture (Lyon et al.,[24]). There were five studies on spinal fusion, including two on anterior lumbar interbody fusion (ALIF) (Burkus et al.[25]; Burkus et al.,[26]) and three on posterior fusion (Haid et al.,[27]; Cho et al.,[28]; Choi et al.,[29]). There were seven studies on maxillofacial bone defects, including four on post-extraction alveolar socket defects (Fiorellini et al.,[30]; Coomes et al.,[31]; Wei et al.,[32] 2024; Jo et al.,[33]) and three on secondary alveolar cleft repair (Dickinson et al.,[34]; Canan et al.,[35]; Alonso et al.,[36]). The follow-up period varied widely, with the shortest being 1.5 months, the longest being 24 months and the median being six months. Spinal fusion studies generally had a longer follow-up time (24 months). The included studies on long bones exclusively focused on tibial fractures. This was not a result of our selection criteria but rather a reflection of the available literature. During our comprehensive search, we identified no RCTs investigating rhBMP-2 for acute fractures of other long bones (e.g., femur, humerus, radius). Detailed literature characteristics are presented in Table I.[9,23-36]

Publication bias assessment

The funnel plot for the primary outcome is presented in Figure 2. Visual inspection showed a usually symmetrical distribution of studies around the pooled effect estimate. The studies by Lyon et al.,[24] Coomes et al.[31] and Wei et al.[32] were observed to lie furthest from the pseudoconfidence interval lines.
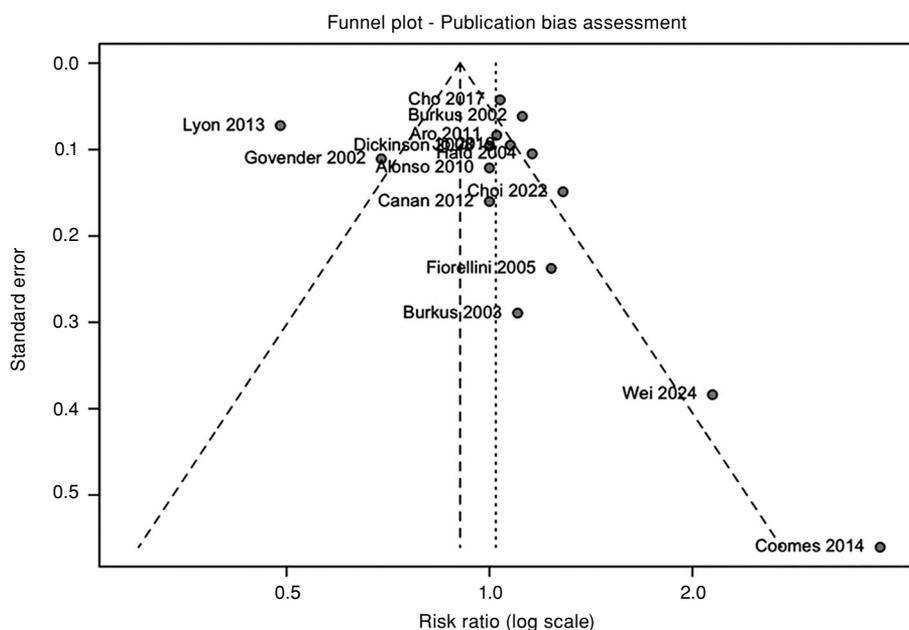
FIGURE 2. Funnel plot publication bias assessment. It assesses publication bias through a funnel plot, with a symmetrical funnel plot indicating no obvious publication bias.

Risk of bias assessment

The 15 RCTs included in this article were all assessed for quality using the Cochrane Collaboration's RoB 2.0 tool, with the results shown in Figure 3. The quality of the included studies was good overall, and most studies exhibited low risk in terms of random sequence generation, selective reporting bias and incomplete outcome data. However, some studies still had unclear risks in terms of allocation concealment (eight articles) and blinding of outcome measurement (seven articles), suggesting that there may be a certain degree of bias risk. All studies showed low risk in terms of incomplete outcome data and selective reporting bias, and the overall data integrity and reporting transparency were high. Therefore, the overall quality of the literature included in this study was good.

Meta-analysis results

This study conducted a meta-analysis of the overall efficacy of the 15 RCTs (Figure 4, Table II). The results showed that there was no significant difference in the overall success rate of bone regeneration in the rhBMP-2 treatment group compared with the control group (RR = 1.02, 95% CI 0.87-1.20). Heterogeneity analysis showed that there was high heterogeneity between the studies ($I^2 = 88.9\%$, $p < 0.0001$), suggesting that the conclusions should be interpreted with caution. In addition, the prediction interval was 0.55-1.89, indicating that rhBMP-2 treatment may have both favorable and unfavorable results in similar studies in the future. Given the known clinical heterogeneity between indications requiring bone regeneration (e.g., the spine) and fracture healing (e.g., long bones), the overall estimate is

FIGURE 3. Bias risk assessment diagram. It assesses the risk of bias in the included studies, using the Cochrane Collaboration Risk of Bias Tool (RoB 2.0).
BMP-2, bone morphogenetic protein-2.

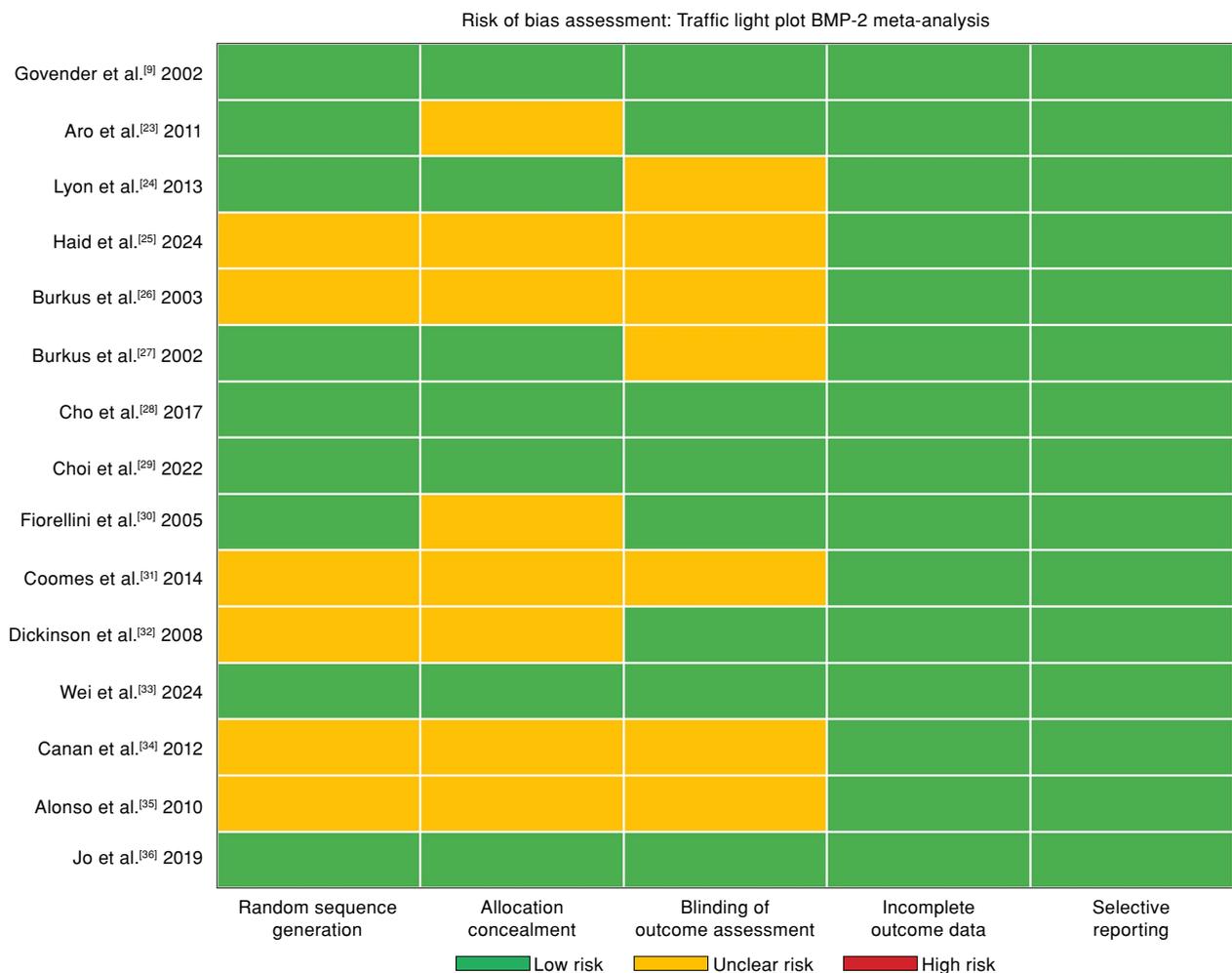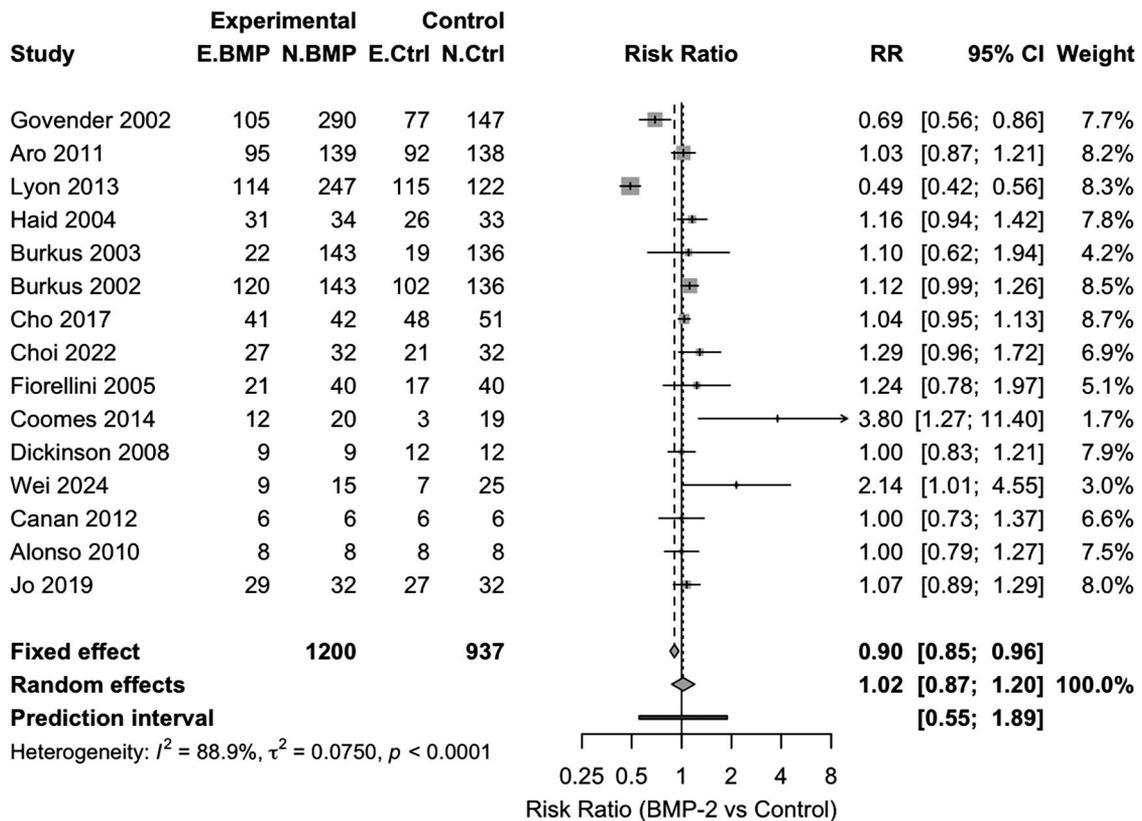| Study	Experimental		Control		Risk Ratio	RR	95% CI	Weight
	E.BMP	N.BMP	E.Ctrl	N.Ctrl				
Govender 2002	105	290	77	147		0.69	[0.56; 0.86]	7.7%
Aro 2011	95	139	92	138		1.03	[0.87; 1.21]	8.2%
Lyon 2013	114	247	115	122		0.49	[0.42; 0.56]	8.3%
Haid 2004	31	34	26	33		1.16	[0.94; 1.42]	7.8%
Burkus 2003	22	143	19	136		1.10	[0.62; 1.94]	4.2%
Burkus 2002	120	143	102	136		1.12	[0.99; 1.26]	8.5%
Cho 2017	41	42	48	51		1.04	[0.95; 1.13]	8.7%
Choi 2022	27	32	21	32		1.29	[0.96; 1.72]	6.9%
Fiorellini 2005	21	40	17	40		1.24	[0.78; 1.97]	5.1%
Coomes 2014	12	20	3	19		3.80	[1.27; 11.40]	1.7%
Dickinson 2008	9	9	12	12		1.00	[0.83; 1.21]	7.9%
Wei 2024	9	15	7	25		2.14	[1.01; 4.55]	3.0%
Canan 2012	6	6	6	6		1.00	[0.73; 1.37]	6.6%
Alonso 2010	8	8	8	8		1.00	[0.79; 1.27]	7.5%
Jo 2019	29	32	27	32		1.07	[0.89; 1.29]	8.0%
Fixed effect		**1200**		**937**		**0.90**	**[0.85; 0.96]**	
Random effects						**1.02**	**[0.87; 1.20]**	**100.0%**
Prediction interval							**[0.55; 1.89]**	

Heterogeneity: $I^2 = 88.9\%$, $\tau^2 = 0.0750$, $p < 0.0001$

Risk Ratio (BMP-2 vs Control)
0.25 0.5 1 2 4 8

FIGURE 4. rhBMP-2 overall forest plot. The pooled RR and 95% CI were calculated using a random-effects model. An RR > 1 indicates a favorable effect for the rhBMP-2 group, suggesting a higher success rate than the control group. The size of the data markers represents the weight of each study in the meta-analysis. Heterogeneity was assessed using the I^2 statistic.
BMP, bone morphogenetic protein; RR, risk ratio; CI, confidence interval.

a composite of different effects and should not be interpreted in isolation. Therefore, we conducted subgroup analyses by carrier type and anatomical site in sequence to provide clinically meaningful insights.

Subgroup analysis was further performed according to the carrier type (Figure 5). No significant difference was observed in the efficacy of the ACS carrier subgroup (RR = 1.03, 95% CI 0.93-1.14), and the heterogeneity was moderate ($I^2 = 55.4\%$). Other synthetic carrier subgroups showed high heterogeneity ($I^2 = 93\%$), and their efficacy CIs were extremely wide (RR = 0.98, 95% CI 0.23-4.14), suggesting that different synthetic carriers may have large differences. The efficacy of the HA carrier subgroup showed a marginal positive effect trend (RR = 1.10, 95% CI 0.91-1.33), but there was no statistically significant difference ($p > 0.05$), and the heterogeneity was low ($I^2 = 48.1\%$).

Considering the analysis according to the anatomical site (Figures 6a-c), the effects of rhBMP-2 showed clear contextual differences. Stomatognathic defects (Figure 6a) showed a positive trend after pooling, but it did not reach statistical significance (random-effects RR = 1.06, 95% CI 0.96-1.18; $I^2 = 38.3\%$; prediction interval 0.93-1.21). Spinal fusion (Figure 6b) showed a consistent and significant benefit (RR = 1.09, 95% CI 1.01-1.17; $I^2 = 0\%$; prediction interval 0.96-1.23). No clear advantage was seen for long bone/tibial fractures (Figure 6c) (RR = 0.70, 95% CI 0.46-1.07), and heterogeneity was very high ($I^2 = 95.5\%$; prediction interval 0.11-4.34), suggesting that differences in injury severity, fracture type, and fixation method/soft tissue conditions across studies may have driven the dispersion of results.

In terms of safety, this study also conducted a meta-analysis of SAEs reported in the included

literature (Figure 7). The results showed that no significant difference was observed in the incidence of SAEs between the rhBMP-2 group and the control group (RR = 0.97, 95% CI 0.66-1.42), and the heterogeneity between studies was low (I^2 = 10.8%), indicating that rhBMP-2 treatment has good overall safety within the scope of the included studies.

Network meta-analysis

The network meta-analysis included the 15 RCTs involving a total of 2,137 patients, and was used to compare the relative efficacy of three treatment options: rhBMP-2 (n = 1,200), autologous bone transplantation (autograft, n = 382) and other bone substitute materials (other, n = 555). The number of research connections showed that the evidence network between the nodes was well connected (Figure 8a). The SUCRA analysis results showed that rhBMP-2 ranked highest in efficacy (SUCRA = 68.3%), followed by other bone substitute materials (SUCRA = 56.3%), with autologous bone transplantation ranking lowest in efficacy (SUCRA = 25.4%) (Figure 8b). This suggests that rhBMP-2 may be superior to the other two treatment methods in promoting bone regeneration.

Using autologous bone as the reference, network meta-analysis (Figure 9) showed that the overall bone regeneration success rate of rhBMP-2 (OR = 1.53, 95% CI 0.54-4.33) and other bone substitute materials (OR = 1.42, 95% CI 0.39-5.21) was not statistically significantly higher than that of autologous bone, but both showed a clear trend of superiority. Autologous bone graft was selected as the reference node for the network meta-analysis as it represents the historical clinical gold standard for bone regeneration.[37] This approach allowed for a direct and clinically interpretable comparison of the relative efficacy of rhBMP-2 and other bone substitute materials against this established benchmark.

Further pairwise direct comparison (rhBMP-2 *vs.* autologous bone) included seven RCTs involving 767 patients (Figure 10). The overall efficacy of rhBMP-2 treatment was significantly better than that of autologous bone transplantation (OR = 1.54, 95% CI 1.02-2.33, $p < 0.05$), and the heterogeneity between studies was low (I^2 = 0%). The direct comparison between rhBMP-2 and other bone substitute materials, presented in Figure 11, was based on data from eight RCTs encompassing 1,370 patients. Here, the

TABLE II

Summary of meta-analysis results

Comparison	Effect measure	Effect size (RR/OR)	95% CI	p	Heterogeneity (I^2) (%)	Prediction interval	Notes
Overall efficacy	RR	1.02	0.87-1.20	0.82	88.90	0.55-1.89	No significant difference, high heterogeneity
Subgroup analysis by carrier type							
ACS carrier	RR	1.03	0.93-1.14	0.62	55.40	-	No significant difference, moderate heterogeneity
Other synthetic carriers	RR	0.98	0.23-4.14	0.98	93.00	-	No significant difference, high heterogeneity
HA carrier	RR	1.1	0.91-1.33	0.3	48.10	-	No significant difference, low heterogeneity
Subgroup analysis by anatomical site							
Oral and maxillofacial bone defects	RR	1.06	0.96-1.18	0.24	0	-	No significant difference, low heterogeneity
Spinal fusion	RR	1.09	1.01-1.17	0.03	0	-	Significant difference, low heterogeneity
Long bone fractures	RR	0.7	0.46-1.07	0.09	95.50	-	No significant difference, high heterogeneity
Safety analysis (SAEs)	RR	0.97	0.66-1.42	0.87	10.80	-	No significant difference, low heterogeneity

RR: Risk ratio; OR: Odds ratio; CI: Confidence interval; ACS: Absorbable collagen sponge; HA: Hydroxyapatite.

Study	Experimental E.BMP	N.BMP	Control E.Ctrl	N.Ctrl	Risk Ratio	RR	95% CI	Weight
Carrier_Category = ACS								
Govender 2002	105	290	77	147		0.69	[0.56; 0.86]	7.7%
Aro 2011	95	139	92	138		1.03	[0.87; 1.21]	8.2%
Haid 2004	31	34	26	33		1.16	[0.94; 1.42]	7.8%
Burkus 2003	22	143	19	136		1.10	[0.62; 1.94]	4.2%
Burkus 2002	120	143	102	136		1.12	[0.99; 1.26]	8.5%
Fiorellini 2005	21	40	17	40		1.24	[0.78; 1.97]	5.1%
Coomes 2014	12	20	3	19		3.80	[1.27; 11.40]	1.7%
Dickinson 2008	9	9	12	12		1.00	[0.83; 1.21]	7.9%
Canan 2012	6	6	6	6		1.00	[0.73; 1.37]	6.6%
Alonso 2010	8	8	8	8		1.00	[0.79; 1.27]	7.5%
Jo 2019	29	32	27	32		1.07	[0.89; 1.29]	8.0%
Fixed effect		864		707		**1.01**	**[0.93; 1.09]**	.
Random effects						**1.03**	**[0.93; 1.14]**	**73.1%**
Heterogeneity: I^2 = 55.4%, τ^2 = 0.0129, p = 0.0130								
Carrier_Category = Other synthetic								
Lyon 2013	114	247	115	122		0.49	[0.42; 0.56]	8.3%
Wei 2024	9	15	7	25		2.14	[1.01; 4.55]	3.0%
Fixed effect		262		147		**0.54**	**[0.47; 0.63]**	.
Random effects						**0.98**	**[0.23; 4.14]**	**11.3%**
Heterogeneity: I^2 = 93%, τ^2 = 1.0134, p = 0.0002								
Carrier_Category = HA								
Cho 2017	41	42	48	51		1.04	[0.95; 1.13]	8.7%
Choi 2022	27	32	21	32		1.29	[0.96; 1.72]	6.9%
Fixed effect		74		83		**1.12**	**[1.00; 1.26]**	.
Random effects						**1.10**	**[0.91; 1.33]**	**15.6%**
Heterogeneity: I^2 = 48.1%, τ^2 = 0.0111, p = 0.1653								
Fixed effect		1200		937		**0.90**	**[0.85; 0.96]**	.
Random effects						**1.02**	**[0.87; 1.20]**	**100.0%**
Prediction interval							**[0.55; 1.89]**	

0.25 0.5 1 2 4 8
Risk Ratio (BMP-2 vs Control)

Heterogeneity: I^2 = 88.9%, τ^2 = 0.0750, p < 0.0001
Test for subgroup differences (common effect): χ^2 = 71.13, df = 2 (p < 0.0001)
Test for subgroup differences (random effect): χ^2 = 0.39, df = 2 (p < 0.8229)

FIGURE 5. Carrier type subgroup forest plot. The pooled RR and 95% CI were calculated using a random-effects model. An RR > 1 indicates a favorable effect for the rhBMP-2 group, suggesting a higher success rate than the control group. The size of the data markers represents the weight of each study in the meta-analysis. Heterogeneity was assessed using the I^2 statistic.
BMP, bone morphogenetic protein; RR, risk ratio; CI, confidence interval.

overall effect of rhBMP-2 treatment did not show a statistically significant difference (OR = 1.17, 95% CI 0.40-3.42), but the heterogeneity was high (I^2 = 90.1%).

Heterogeneity exploration based on machine learning

This study further used ML methods to explore the sources of heterogeneity in the meta-analysis.
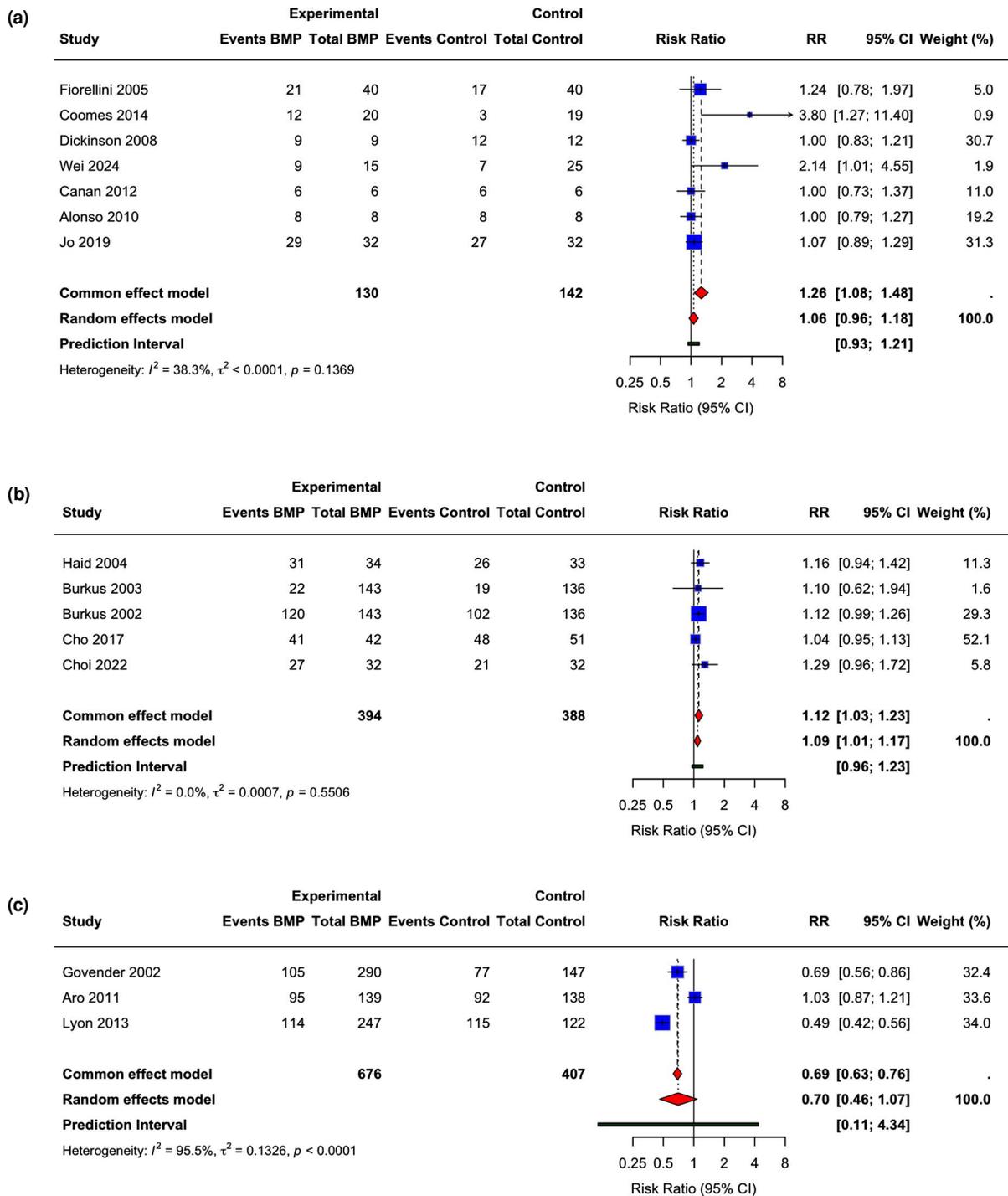
(a)

Study	Experimental		Control		Risk Ratio	RR	95% CI	Weight (%)
	Events BMP	Total BMP	Events Control	Total Control				
Fiorellini 2005	21	40	17	40		1.24	[0.78; 1.97]	5.0
Coomes 2014	12	20	3	19		3.80	[1.27; 11.40]	0.9
Dickinson 2008	9	9	12	12		1.00	[0.83; 1.21]	30.7
Wei 2024	9	15	7	25		2.14	[1.01; 4.55]	1.9
Canan 2012	6	6	6	6		1.00	[0.73; 1.37]	11.0
Alonso 2010	8	8	8	8		1.00	[0.79; 1.27]	19.2
Jo 2019	29	32	27	32		1.07	[0.89; 1.29]	31.3
Common effect model		**130**		**142**		**1.26**	**[1.08; 1.48]**	.
Random effects model						**1.06**	**[0.96; 1.18]**	100.0
Prediction Interval							**[0.93; 1.21]**	

Heterogeneity: $I^2 = 38.3\%$, $\tau^2 < 0.0001$, $p = 0.1369$

0.25 0.5 1 2 4 8
Risk Ratio (95% CI)

(b)

Study	Experimental		Control		Risk Ratio	RR	95% CI	Weight (%)
	Events BMP	Total BMP	Events Control	Total Control				
Haid 2004	31	34	26	33		1.16	[0.94; 1.42]	11.3
Burkus 2003	22	143	19	136		1.10	[0.62; 1.94]	1.6
Burkus 2002	120	143	102	136		1.12	[0.99; 1.26]	29.3
Cho 2017	41	42	48	51		1.04	[0.95; 1.13]	52.1
Choi 2022	27	32	21	32		1.29	[0.96; 1.72]	5.8
Common effect model		**394**		**388**		**1.12**	**[1.03; 1.23]**	.
Random effects model						**1.09**	**[1.01; 1.17]**	100.0
Prediction Interval							**[0.96; 1.23]**	

Heterogeneity: $I^2 = 0.0\%$, $\tau^2 = 0.0007$, $p = 0.5506$

0.25 0.5 1 2 4 8
Risk Ratio (95% CI)

(c)

Study	Experimental		Control		Risk Ratio	RR	95% CI	Weight (%)
	Events BMP	Total BMP	Events Control	Total Control				
Govender 2002	105	290	77	147		0.69	[0.56; 0.86]	32.4
Aro 2011	95	139	92	138		1.03	[0.87; 1.21]	33.6
Lyon 2013	114	247	115	122		0.49	[0.42; 0.56]	34.0
Common effect model		**676**		**407**		**0.69**	**[0.63; 0.76]**	.
Random effects model						**0.70**	**[0.46; 1.07]**	100.0
Prediction Interval							**[0.11; 4.34]**	

Heterogeneity: $I^2 = 95.5\%$, $\tau^2 = 0.1326$, $p < 0.0001$

0.25 0.5 1 2 4 8
Risk Ratio (95% CI)

FIGURE 6. Anatomical site subgroup forest plot. **(a)** Site category= Oral maxillofacial; **(b)** Site category= Spine; **(c)** Site Category = Long bone. The pooled RR and 95% CI were calculated using a random-effects model. An RR > 1 indicates a favorable effect for the rhBMP-2 group, suggesting a higher success rate than the control group. The size of the data markers represents the weight of each study in the meta-analysis. Heterogeneity was assessed using the I^2 statistic.
BMP, bone morphogenetic protein; RR, risk ratio; CI, confidence interval.

The feature importance analysis of the random forest and gradient boosting models (Figure 12) showed that the main sources of heterogeneity in the meta-analysis were the total sample size and sample size ratio of the study, which were significantly more important than other features. In addition, the
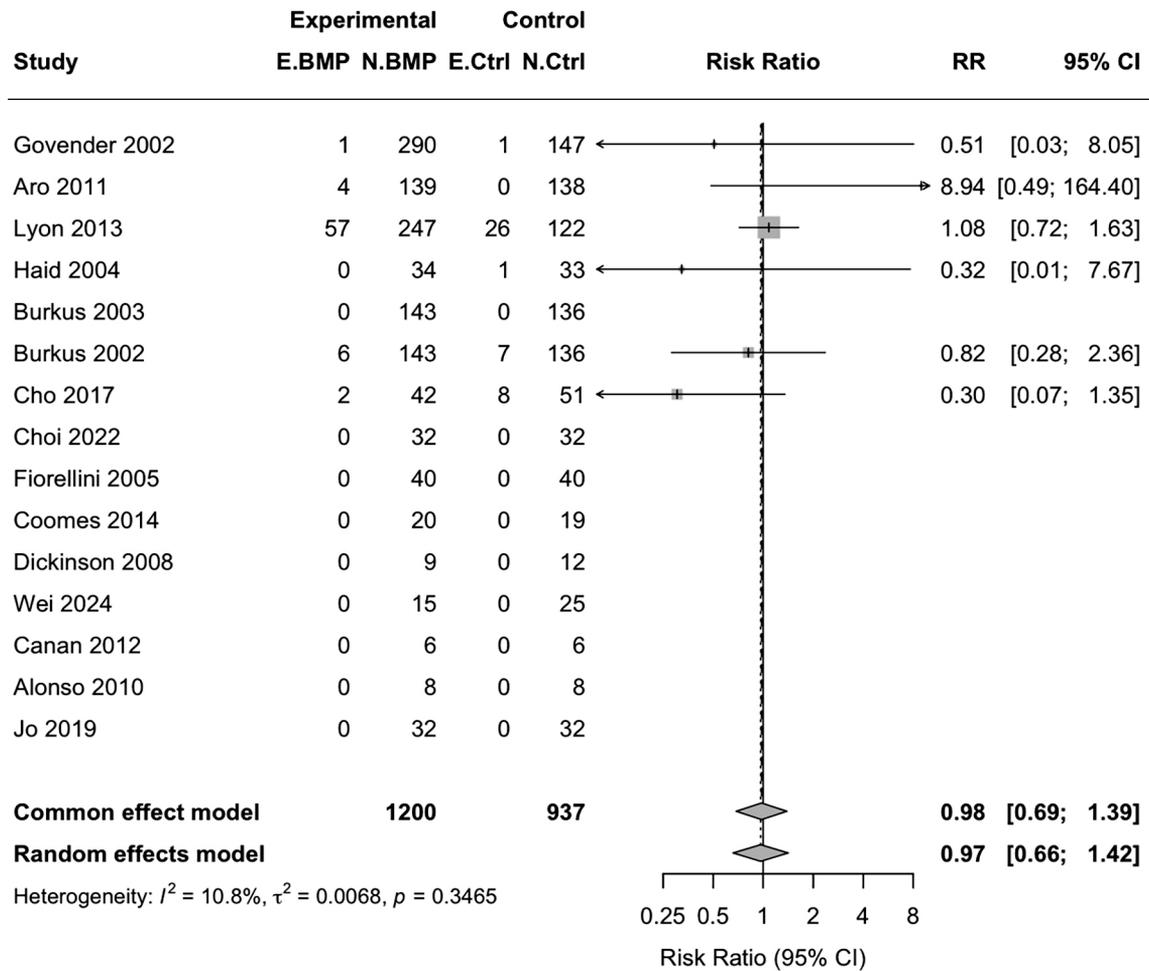
Study	Experimental		Control		Risk Ratio	RR	95% CI
	E.BMP	N.BMP	E.Ctrl	N.Ctrl			
Govender 2002	1	290	1	147		0.51	[0.03; 8.05]
Aro 2011	4	139	0	138		8.94	[0.49; 164.40]
Lyon 2013	57	247	26	122		1.08	[0.72; 1.63]
Haid 2004	0	34	1	33		0.32	[0.01; 7.67]
Burkus 2003	0	143	0	136			
Burkus 2002	6	143	7	136		0.82	[0.28; 2.36]
Cho 2017	2	42	8	51		0.30	[0.07; 1.35]
Choi 2022	0	32	0	32			
Fiorellini 2005	0	40	0	40			
Coomes 2014	0	20	0	19			
Dickinson 2008	0	9	0	12			
Wei 2024	0	15	0	25			
Canan 2012	0	6	0	6			
Alonso 2010	0	8	0	8			
Jo 2019	0	32	0	32			
Common effect model		**1200**		**937**		**0.98**	**[0.69; 1.39]**
Random effects model						**0.97**	**[0.66; 1.42]**

Heterogeneity: $I^2 = 10.8\%$, $\tau^2 = 0.0068$, $p = 0.3465$

Risk Ratio (95% CI): 0.25 0.5 1 2 4 8

FIGURE 7. Serious adverse events forest plot. The pooled RR and 95% CI were calculated using a random-effects model. An RR > 1 indicates a favorable effect for the rhBMP-2 group, suggesting a higher success rate than the control group. The size of the data markers represents the weight of each study in the meta-analysis. Heterogeneity was assessed using the I^2 statistic.
BMP, bone morphogenetic protein; RR, risk ratio; CI, confidence interval.

mean age (year) of the subjects, the year of study publication (year) and the total dose of rhBMP-2 (dose total mg) were also identified by the model as secondary important features, while the carrier type (carrier) and anatomical site (site) contributed relatively little to the importance of the model.

The distribution of the effect size (RR) and subgroup characteristics of each study were further analyzed (Figure 13). Both RR and Log-RR showed a relatively concentrated distribution, but some individual studies still showed obvious deviations in effect size, suggesting the existence of potential extreme effect values. Subgroup analysis based on the anatomical site of the study showed that the alveolar socket group

had a more obvious positive therapeutic effect trend, while the closed tibial fracture subgroup (tibia-closed group) showed a negative effect trend, suggesting that clinical indications may be one of the important sources of effect heterogeneity.

DISCUSSION

It is of utmost importance to clearly distinguish between the concepts of bone healing and bone regeneration. Bone healing primarily involves the process of fracture site union, which usually refers to the formation and remodeling of callus between the fracture ends to achieve bony connection. This process is influenced by fracture type, severity of injury and micromotion under physiological
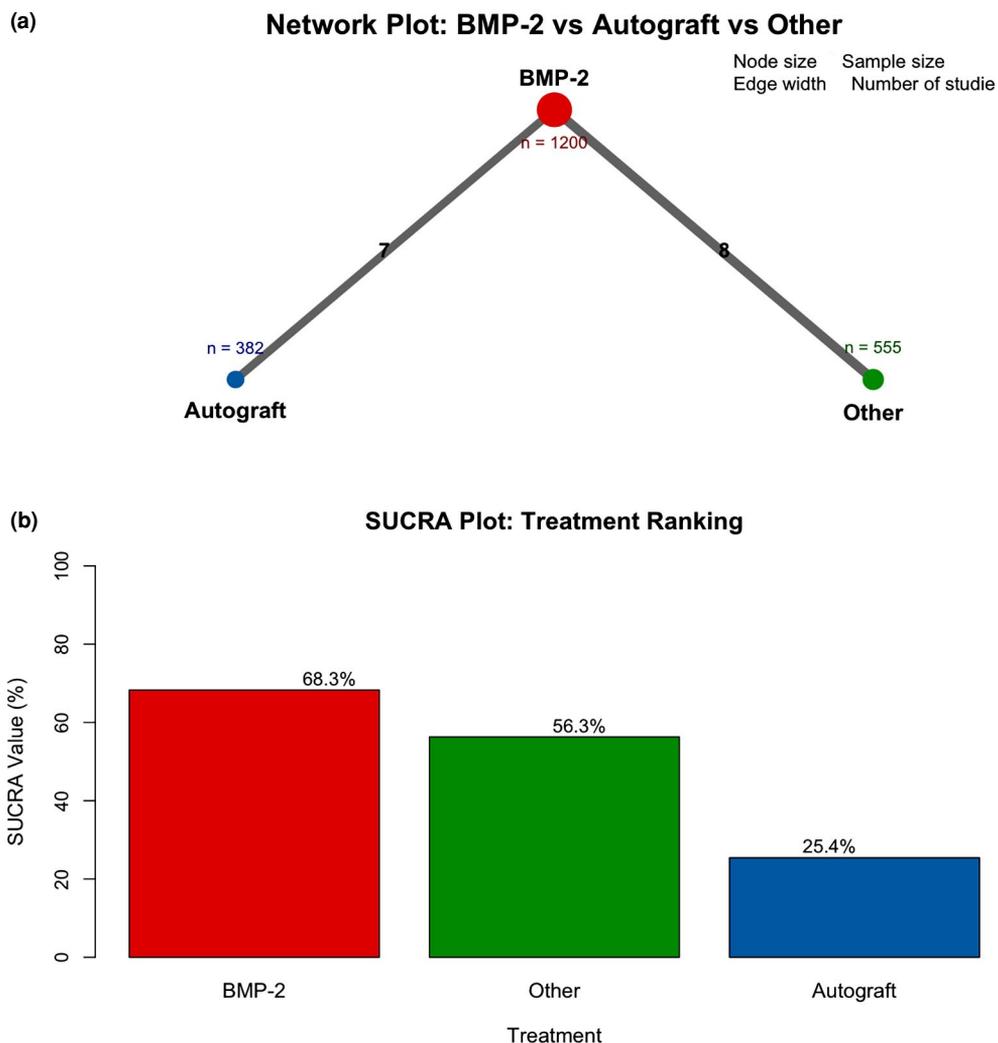
(a)

Network Plot: BMP-2 vs Autograft vs Other

Node size   Sample size
Edge width   Number of studie

BMP-2

n = 1200

7   8

n = 382   n = 555

Autograft   **Other**

(b)   **SUCRA Plot: Treatment Ranking**

68.3%

56.3%

25.4%

BMP-2   Other   Autograft

Treatment

FIGURE 8. Network Meta-analysis. **(a)** Network evidence relationship diagram. It shows the network of evidence relationships between different treatment nodes. **(b)** SUCRA ranking probability diagram. It shows the efficacy ranking probability of different treatment options.
BMP, bone morphogenetic protein; SUCRA, surface under the cumulative ranking curve.

conditions. Bone regeneration, on the other hand, mainly involves the repair of bone defects, which usually refers to the formation of new bone to fill

Comparison: other vs 'Autograft'

Treatment	(Random Effects Model)	OR	95%-CI
BMP-2		1.53	[0.54; 4.33]
Other		1.42	[0.39; 5.21]

0.25 0.5   1   2   4   8
Odds Ratio vs Autograft

FIGURE 9. Network Meta effect synthesis diagram. It shows the synthesized efficacy results of BMP-2, autograft, and other bone substitute materials relative to autograft.
BMP, bone morphogenetic protein; OR, odds ratio; CI, confidence interval.
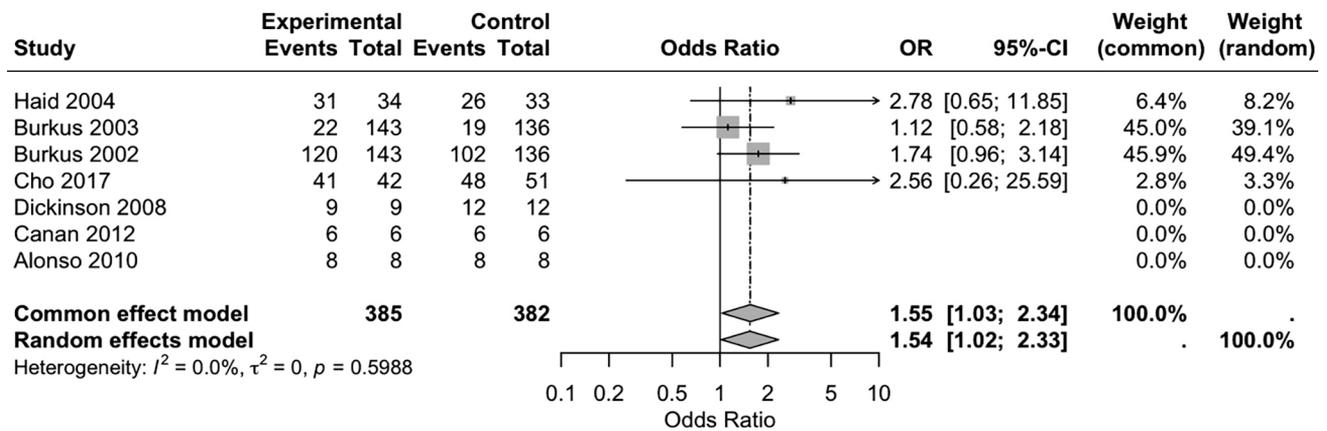
the defect, such as alveolar socket defects and palatal cleft repair. This study comprehensively evaluated the efficacy and safety of rhBMP-2 in promoting bone healing and regeneration through a systematic review based on ML, meta-analysis, network meta-analysis and heterogeneity analysis. The meta-analysis results showed that the overall efficacy in the rhBMP-2 group was not significantly different from that of the control group (RR = 1.02, 95% CI 0.87-1.20). The subgroup analysis by anatomical site for the oral and maxillofacial bone defect subgroup, spinal fusion subgroup and long bone fracture subgroup revealed that only the spinal fusion subgroup had stable and significant efficacy in promoting bone healing (RR = 1.09, 95% CI

Study	Experimental Events	Total	Control Events	Total	Odds Ratio	OR	95%-CI	Weight (common)	Weight (random)
Haid 2004	31	34	26	33		2.78	[0.65; 11.85]	6.4%	8.2%
Burkus 2003	22	143	19	136		1.12	[0.58; 2.18]	45.0%	39.1%
Burkus 2002	120	143	102	136		1.74	[0.96; 3.14]	45.9%	49.4%
Cho 2017	41	42	48	51		2.56	[0.26; 25.59]	2.8%	3.3%
Dickinson 2008	9	9	12	12				0.0%	0.0%
Canan 2012	6	6	6	6				0.0%	0.0%
Alonso 2010	8	8	8	8				0.0%	0.0%
Common effect model		**385**		**382**		**1.55**	**[1.03; 2.34]**	**100.0%**	**.**
Random effects model						**1.54**	**[1.02; 2.33]**	**.**	**100.0%**

Heterogeneity: $I^2 = 0.0\%$, $\tau^2 = 0$, $p = 0.5988$

0.1 0.2   0.5 1 2   5 10
Odds Ratio

FIGURE 10. rhBMP-2 *vs.* autologous bone forest plot. The pooled RR and 95% CI were calculated using a random-effects model. An RR > 1 indicates a favorable effect for the rhBMP-2 group, suggesting a higher success rate than the control group. The size of the data markers represents the weight of each study in the meta-analysis. Heterogeneity was assessed using the I^2 statistic. OR, odds ratio; CI, confidence interval.

Study	Experimental Events	Total	Control Events	Total	Odds Ratio	OR	95%-CI	Weight (common)	Weight (random)
Govender 2002	105	290	77	147		0.52	[0.34; 0.77]	33.5%	14.0%
Aro 2011	95	139	92	138		1.08	[0.65; 1.79]	15.0%	13.8%
Lyon 2013	114	247	115	122		0.05	[0.02; 0.12]	42.6%	13.2%
Choi 2022	27	32	21	32		2.83	[0.85; 9.40]	1.7%	12.1%
Fiorellini 2005	21	40	17	40		1.50	[0.62; 3.61]	4.2%	13.0%
Coomes 2014	12	20	3	19		8.00	[1.74; 36.70]	0.6%	11.1%
Wei 2024	9	15	7	25		3.86	[1.00; 14.92]	1.1%	11.6%
Jo 2019	29	32	27	32		1.79	[0.39; 8.22]	1.3%	11.1%
Common effect model		**815**		**555**		**0.58**	**[0.46; 0.73]**	**100.0%**	**.**
Random effects model						**1.17**	**[0.40; 3.42]**	**.**	**100.0%**

Heterogeneity: $I^2 = 90.1\%$, $\tau^2 = 2.0970$, $p < 0.0001$

0.1 0.2   0.5 1 2   5 10
Odds Ratio
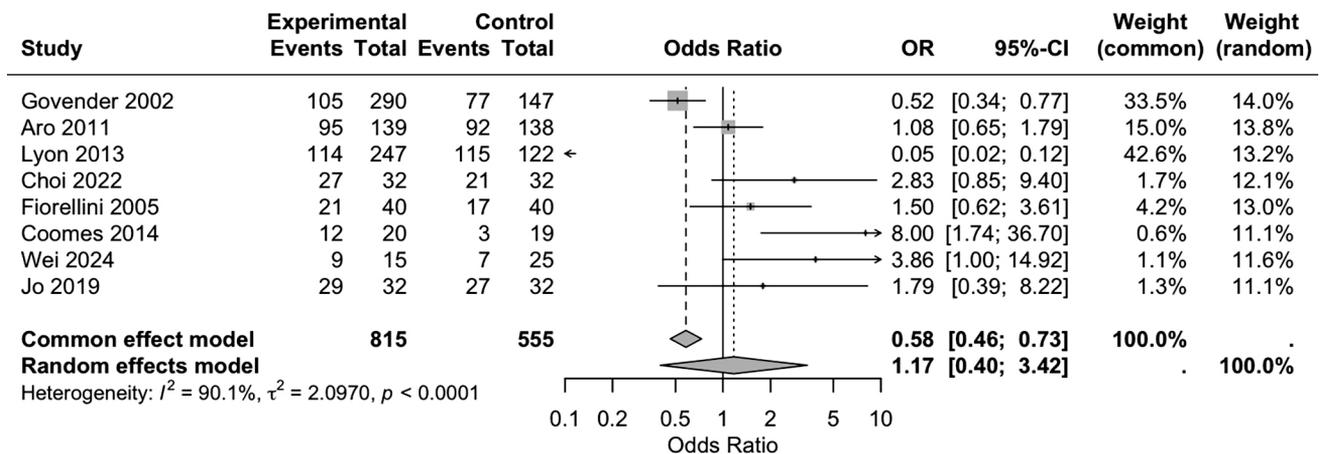
FIGURE 11. rhBMP-2 *vs.* other materials forest plot. The pooled RR and 95% CI were calculated using a random-effects model. An RR > 1 indicates a favorable effect for the rhBMP-2 group, suggesting a higher success rate than the control group. The size of the data markers represents the weight of each study in the meta-analysis. Heterogeneity was assessed using the I^2 statistic. OR, odds ratio; CI, confidence interval.

1.01-1.17, $p < 0.05$), with low heterogeneity ($I^2 = 0\%$), which may be related to the severity of injury and fracture type of the patients.

Network meta-analysis further confirmed that the efficacy of rhBMP-2 treatment had a clear advantage over autologous bone treatment, whereas there was no significant difference compared to other synthetic bone substitute materials, but there was high heterogeneity. Safety analysis showed that rhBMP-2 treatment had good overall safety and did not increase the risk of SAEs. The ML model revealed that the effect heterogeneity in the

meta-analysis was mainly due to sample size, age and dose factors. Based on the above results, this study further supports the application prospects of rhBMP-2 in bone regeneration, and we believe that differentiated application strategies for specific populations and specific clinical scenarios should be considered in clinical applications.

Currently, clinical trials of rhBMP-2, as a potent osteoinductive factor, have been widely reported.[38-40] Overall, this study did not find a significant advantage of rhBMP-2 over controls, a result similar to the results of several recently
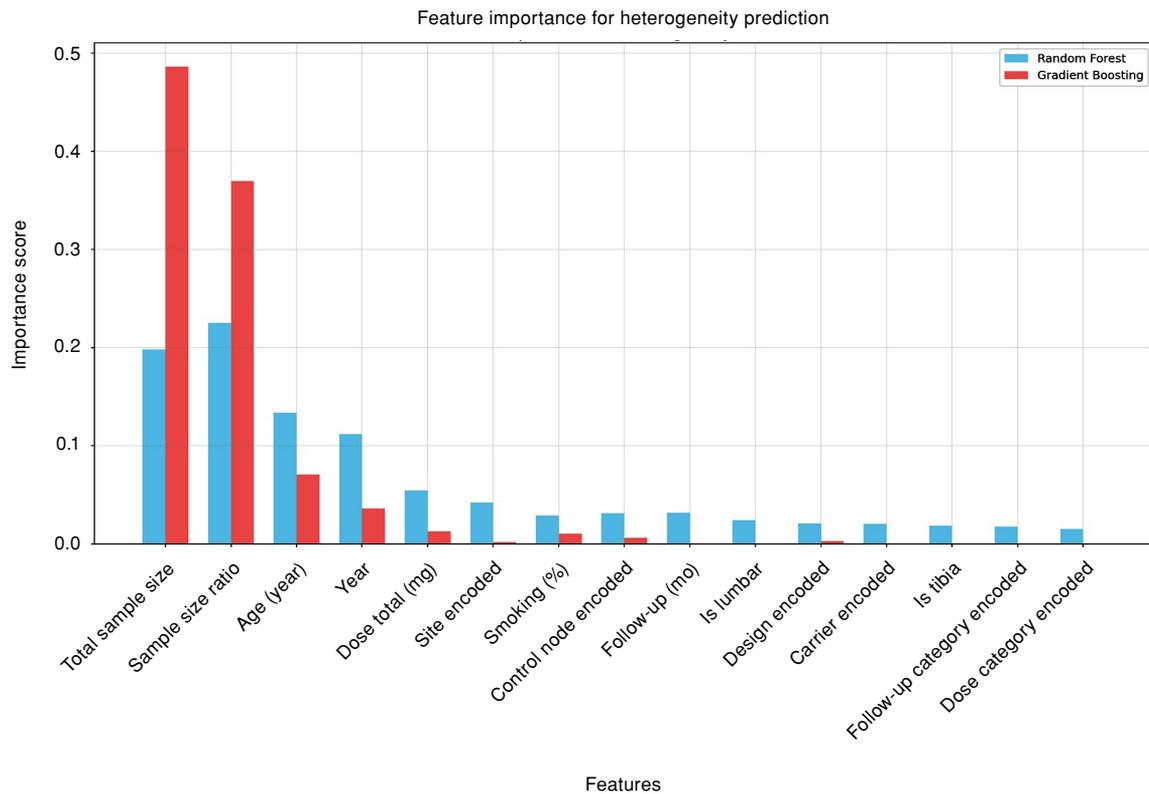
Feature importance for heterogeneity prediction

FIGURE 12. Heterogeneity analysis of feature importance based on random forest and gradient boosting models.

published systematic reviews.[41-43] A recent meta-analysis of rhBMP-2 in spinal fusion found that rhBMP-2 could effectively improve the fusion rate, but its advantage was not stable in small-sample studies.[17,44] Similarly, the present study also found significant heterogeneity, which suggests that the diversity of study designs may mask the true treatment effect.

Further subgroup analysis results showed that rhBMP-2 showed a clear efficacy advantage in spinal fusion (RR = 1.09), which is consistent with the efficacy advantage reported by Burkus et al.[25,26] in ALIF surgery. However, no significant effect was observed in the study of long bone fractures, which may be related to the fact that the healing process of such fractures is affected by multiple factors, including fracture type and severity of soft tissue injury. Studies have shown that the healing process of open fractures is more complicated, and the use of rhBMP-2 alone cannot ensure good healing.[9,14]

Network meta-analysis showed that rhBMP-2 treatment showed significant advantages over autologous bone treatment (OR = 1.54), with this result verified by several recent RCT studies. Autologous bone treatment has been widely used as the gold standard, but it involves certain problems, such as complications in the bone donor area and limited bone supply. As a bioactive material, rhBMP-2 overcomes the limitations associated with autologous bone harvesting, thereby accounting for its therapeutic advantage.[45,46] However, no significant difference was observed, when rhBMP-2 was compared with other bone substitute materials. This may be attributed to advancements in synthetic material technologies, such as hydroxyapatite (HA) and β-tricalcium phosphate (β-TCP) composite carriers, which have also demonstrated favorable clinical outcomes.[32] Therefore, a finding that requires cautious interpretation has emerged: rhBMP-2 is significantly superior to autograft bone, but shows no statistical difference compared with the 'Other' category of bone substitute materials. This is not necessarily contradictory, but rather reflects the nature of the comparison. The superiority over autograft bone is a key finding, demonstrating that rhBMP-2 can provide a more effective osteoinductive stimulus while avoiding
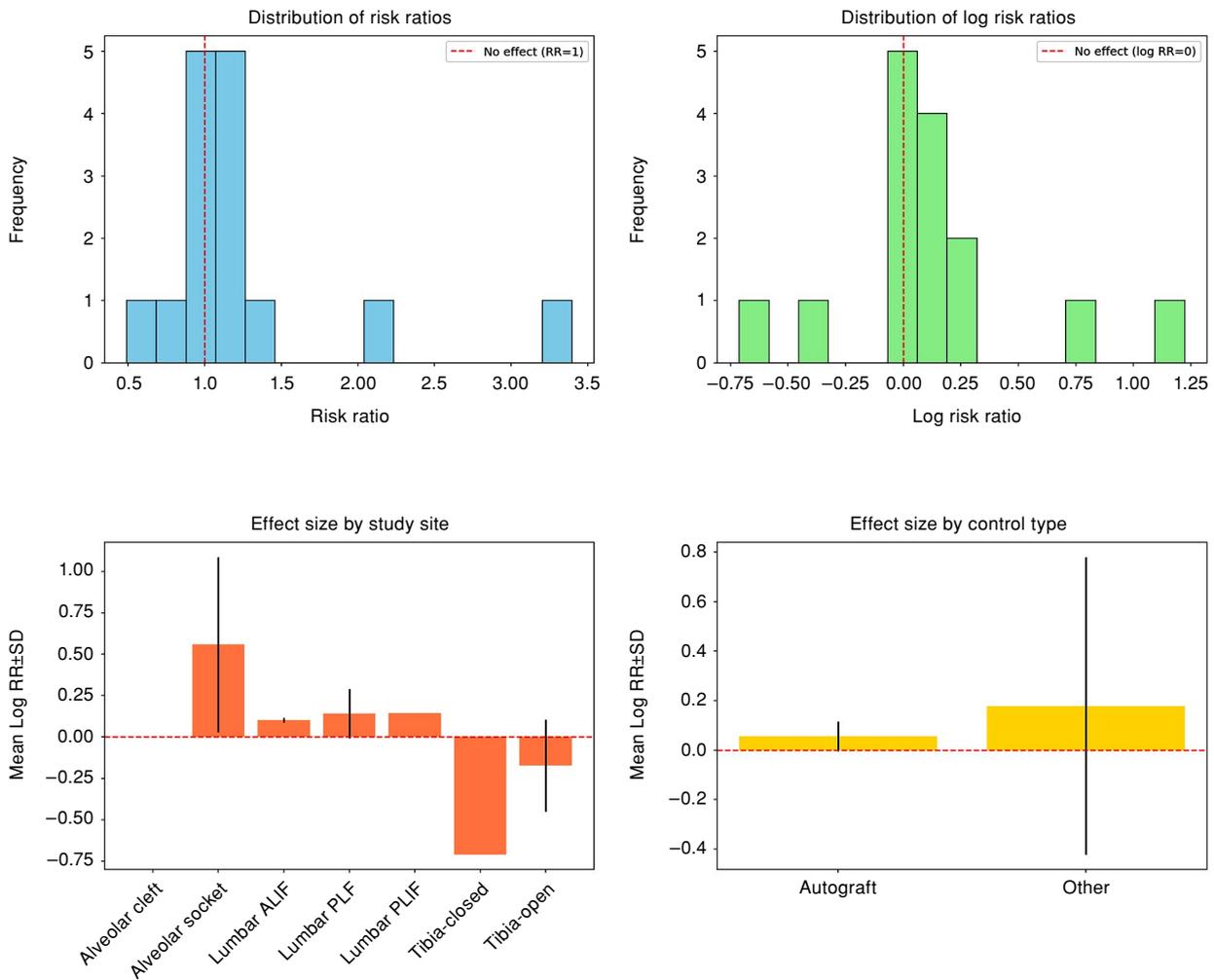
FIGURE 13. Effect size and subgroup distribution diagram. It shows the distribution of effect sizes RR and their Log-RR for each study across different subgroups.
ALIF, anterior lumbar interbody fusion; PLF, posterolateral fusion; PLIF, posterior lumbar interbody fusion; RR, risk ratio.

the inherent donor-site complications associated with autograft bone harvesting. The lack of significant difference compared to the 'Other' group may stem from two factors. First, the 'Other' group consisted of a heterogeneous mixture, including advanced, highly engineered synthetic scaffolds (e.g., HA, β-TCP), which offer excellent osteoconductivity. In certain clinical scenarios, a robust osteoconductive scaffold may be sufficient, which diminishes the measurable additive effect of rhBMP-2's osteoinductive action on a broad average level. Second, the comparison may have lacked the power to detect smaller but clinically relevant differences between rhBMP-2 and these effective synthetic materials. Therefore, this result should not be interpreted as equivalence; rather it

should be understood in terms of both rhBMP-2 and modern synthetic alternatives being viable and often superior alternatives to autograft bone, with the specific choice depending on the clinical need for osteoinduction versus osteoconduction.

Moreover, the statistical superiority of rhBMP-2 over autografts also needs to be interpreted within its clinical and economic context. The primary clinical benefit of rhBMP-2 lies in its ability to circumvent the donor-site morbidity, including persistent pain, infection and neurovascular injury, associated with autograft harvesting. This makes it a particularly valuable option for patients with limited autologous bone availability, those requiring large-volume reconstruction or for those for whom

minimizing operative time and complexity is a priority. However, the high cost of rhBMP-2 biomaterials remains a significant constraint on its widespread use. Therefore, the cost-benefit ratio may be most favorable in specific scenarios, such as complex spinal fusions or maxillofacial reconstructions, where the avoidance of a secondary donor-site surgery justifies the initial expense. Future health-economic studies are needed to formally evaluate the cost-effectiveness of rhBMP-2 across different surgical indications and healthcare systems. Ultimately, patient selection should be guided by a holistic consideration of the defect characteristics, patient comorbidities, surgical expertise and economic factors, rather than efficacy alone.

The subgroup analysis revealed a marginal positive effect trend for HA carriers. This trend may be attributed to the inherent advantages of HA as a biomaterial, rather than the effect of rhBMP-2 alone; HA closely mimics the inorganic component of bone matrix, providing an exceptionally osteoconductive scaffold that promotes excellent osseointegration and new bone formation. Furthermore, the physical and chemical properties of HA carriers may facilitate more favorable release kinetics for rhBMP-2, potentially offering a more sustained release profile compared with rapid-release carriers such as collagen sponges, thereby maintaining an effective therapeutic concentration at the implant site for a longer duration. This suggests a potential synergistic effect, where the osteoconductive properties of the HA scaffold enhance the osteoinductive efficacy of rhBMP-2. Future research should directly compare different carrier systems with a focus on their release kinetics and subsequent bone regeneration outcomes to optimize carrier selection for specific clinical applications.

Regarding the mechanism of this study, we believe that the mechanism of rhBMP-2 promoting bone regeneration involves inducing the chemotaxis and differentiation of mesenchymal stem cells and the deposition of bone matrix. The carrier material plays an important role in this process, controlling the release kinetics of rhBMP-2 and stabilizing its biological activity, thus affecting the stability of the therapeutic effect.[47] As a result, the heterogeneity of different carrier types is significant, which suggests that the carrier material may directly determine the local release pattern and activity retention of rhBMP-2, further affecting the final clinical effect.

In terms of confounding factors, factors such as age, sample size and dose were identified by the ML model as important sources of heterogeneity, which also reflects the actual situation in clinical research; that is, the bone regeneration ability of patients of different age groups varies greatly.[20] Due to osteoporosis, insufficient blood supply and the presence of local inflammatory environment, the bone regeneration effect induced by rhBMP-2 in elderly patients may be significantly lower than that in young patients, which may lead to the outstanding contribution of age factors in heterogeneity analysis.

Nonetheless, this study also has some limitations. First, although the number of included studies is relatively sufficient, the differences in dose, follow-up time and efficacy evaluation methods among different studies led to greater heterogeneity, which may have affected the robustness of the results. Second, the incomplete reporting of important confounding factors such as diabetes and smoking may also have affected the research results. Future research is recommended to focus on the following aspects. First, the research design should be strengthened, the follow-up time and efficacy evaluation indicators unified and the comparability between studies increased. Second, the reporting and analysis of patients' metabolic factors (e.g., diabetes, osteoporosis) and behavioral factors (e.g., smoking) should be strengthened. Third, the mechanism of carrier materials on rhBMP-2 release and activity retention should be explored, and finally, real-world evidence research should be conducted to improve the external validity of research results.

In conclusion, this study comprehensively evaluated the effectiveness and safety of rhBMP-2 for bone regeneration. The results indicate that while rhBMP-2 does not demonstrate a significant overall advantage, it shows a potent and clinically valuable effect in specific indications, particularly spinal fusion and small maxillofacial bone defects, where its efficacy surpasses that of autograft bone without the associated donor-site morbidity. Based on our findings, we recommend the selective use of rhBMP-2 in these well-defined clinical scenarios. Its application in complex long bone fractures; however, it should be exercised with caution and warrants further investigation. The heterogeneity in treatment effects underscores the importance of patient-specific factors, such as age and defect characteristics, as well as technical considerations such as carrier selection and dosing. Future high-quality RCTs and mechanistic studies are

crucial to optimize rhBMP-2 application strategies, refine carrier systems and establish cost-effective protocols for its use in clinical practice.

Data Sharing Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Author Contributions: L.Y.: Conceptualization; L.Y., J.Z., C.C.: Data curation, formal analysis, investigation, methodology, resources, software, writing-original draft, writing-review & editing; L.Y., J.Z.: Supervision; L.Y., C.C.: Validation, visualization. All authors read and approved the final manuscript.

Conflict of Interest: The authors declared no conflicts of interest with respect to the authorship and/or publication of this article.

Funding: The authors received no financial support for the research and/or authorship of this article.

AI Disclosure: The authors declare that artificial intelligence (AI) tools were not used, or were used solely for language editing, and had no role in data analysis, interpretation, or the formulation of conclusions. All scientific content, data interpretation, and conclusions are the sole responsibility of the authors. The authors further confirm that AI tools were not used to generate, fabricate, or 'hallucinate' references, and that all references have been carefully verified for accuracy.

REFERENCES

1. Echave MC, Erezuma I, Golafshan N, Castilho M, Kadumudi FB, Pimenta-Lopes C, et al. Bioinspired gelatin/bioceramic composites loaded with Bone Morphogenetic Protein-2 (BMP-2) promote osteoporotic bone repair. Biomater Adv 2022;134:112539. doi: 10.1016/j.msec.2021.112539.

2. Fu Z, Li D, Lin K, Zhao B, Wang X. Enhancing the osteogenic differentiation of aligned electrospun poly(L-lactic acid) nanofiber scaffolds by incorporation of bioactive calcium silicate nanowires. Int J Biol Macromol 2023;226:1079-87. doi: 10.1016/j.ijbiomac.2022.11.224.

3. Bo R, Yongliang T, Ni L. Bangding, thermosensitive antibacterial hydrogel for treatment of infected bone defects. CJTER 2025;29:7269-77.

4. Aghaloo T, Valentini P, Yardley R, Zadeh HH. Application of biologics in maxillary sinus augmentation surgery: A narrative review. Clin Implant Dent Relat Res 2025;27:e70004. doi: 10.1111/cid.70004.

5. Xufeng J, Miao L, Guangping H. Calcium phosphate combined with recombinant human bone morphogenetic protein-2 in repair and reconstruction of tibial infectious bone defects. CJTER 2024;28:2625-30.

6. Reddy SSP, Francis DL, Thirumoorthi H, Rahul, Rathi M, Singh H, et al. Effectiveness of recombinant human Bone Morphogenetic Protein-2 in socket preservation: A randomized controlled clinical and sequential human histological trial (BMP-2 TRIAL). Clin Exp Dent Res 2025;11:e70134. doi: 10.1002/cre2.70134.

7. Marshall KM, McLaren JS, Wojciechowski JP, Callens SJP, Echalier C, Kanczler JM, et al. Bioactive coatings on 3D printed scaffolds for bone regeneration: Use of Laponite® to deliver BMP-2 in an ovine femoral condyle defect model. Biomater Adv 2024;164:213959. doi: 10.1016/j.bioadv.2024.213959.

8. Briquez P, Tsai HM, Watkins E, Hubbell J. Engineering of a bridge protein to improve the delivery of BMP-2 from collagen sponge and enhance bone regeneration for spinal fusion. Tissue Engineering - Part A 2023;29:670-1.

9. Govender S, Csimma C, Genant HK, Valentin-Opran A, Amit Y, Arbel R, et al. Recombinant human bone morphogenetic protein-2 for treatment of open tibial fractures: A prospective, controlled, randomized study of four hundred and fifty patients. J Bone Joint Surg Am 2002;84:2123-34. doi: 10.2106/00004623-200212000-00001.

10. Xu H, Liao H, Liu X, Miller AL 2nd, Elder BD, Lu L. Spinal fusion of biodegradable poly(propylene fumarate) and poly(propylene fumarate-co-caprolactone) copolymers in rabbits. J Orthop 2023;48:52-9. doi: 10.1016/j.jor.2023.10.023.

11. Sandhu H. Spinal fusion using bone morphogenetic proteins. Orthopedics 2004;27:717-8. doi: 10.3928/0147-7447-20040701-10.

12. Liao SS, Guan K, Cui FZ, Shi SS, Sun TS. Lumbar spinal fusion with a mineralized collagen matrix and rhBMP-2 in a rabbit model. Spine (Phila Pa 1976) 2003;28:1954-60. doi: 10.1097/01.BRS.0000083240.13332.F6.

13. Wang R, Liu W, Guo H, Ge S, Huang H, Yang P. Alveolar ridge preservation with fibroblast growth factor-2 modified acellular dermal matrix membrane and a bovine-derived xenograft: An experimental in vivo study. Clin Oral Implants Res 2021;32:808-17. doi: 10.1111/clr.13749.

14. Alawami EAA, Alomari F, Aloqaybi SA, Aloweiny Q, Alswayed LK, Alshafai NW, et al. Efficacy of recombinant human bone morphogenetic protein-2 in alveolar cleft treatment for children: Systematic review and meta-analysis. Life (Basel) 2025;15:185. doi: 10.3390/life15020185.

15. Tian X, Vater C, Raina DB, Findeisen L, Matuszewski LM, Tägil M, et al. Co-delivery of rhBMP-2 and zoledronic acid using calcium sulfate/hydroxyapatite carrier as a bioactive bone substitute to enhance and accelerate spinal fusion. Bioact Mater 2024;36:256-71. doi: 10.1016/j.bioactmat.2024.02.034.

16. Han SH, Cha M, Jin YZ, Lee KM, Lee JH. BMP-2 and hMSC dual delivery onto 3D printed PLA-Biogel scaffold for critical-size bone defect regeneration in rabbit tibia. Biomed Mater 2020;16:015019. doi: 10.1088/1748-605X/aba879.

17. Yang S, Zhou B, Mo J, He R, Mei K, Zeng Z, et al. Risk factors affecting spinal fusion: A meta-analysis of 39 cohort studies. PLoS One 2024;19:e0304473. doi: 10.1371/journal.pone.0304473.

18. Wen YD, Jiang WM, Yang HL, Shi JH. Exploratory meta-analysis on dose-related efficacy and complications of rhBMP-2 in anterior cervical discectomy and fusion: 1,539,021 Cases from 2003 to 2017 studies. J Orthop Translat 2020;24:166-74. doi: 10.1016/j.jot.2020.01.002.

19. Schoonraad SA, Jaimes AA, Singh AJX, Croland KJ, Bryant SJ. Osteogenic effects of covalently tethered rhBMP-2 and rhBMP-9 in an MMP-sensitive PEG hydrogel nanocomposite. Acta Biomater 2023;170:53-67. doi: 10.1016/j.actbio.2023.08.045.

20. Parums DV. Editorial: Review Articles, Systematic Reviews, Meta-Analysis, and the Updated Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 Guidelines. Med Sci Monit 2021;27:e934475. doi: 10.12659/MSM.934475.

21. Dąbrowska-Bender M, Kozaczuk A, Pączek L, Milkiewicz P, Słoniewski R, Staniszewska A. Patient quality of life after liver transplantation in terms of emotional problems and the impact of sociodemographic factors. Transplant Proc 2018;50:2031-8. doi: 10.1016/j.transproceed.2018.03.113.

22. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928. doi: 10.1136/bmj.d5928.

23. Aro HT, Govender S, Patel AD, Hernigou P, Perera de Gregorio A, Popescu GI, et al. Recombinant human bone morphogenetic protein-2: A randomized trial in open tibial fractures treated with reamed nail fixation. J Bone Joint Surg Am 2011;93:801-8. doi: 10.2106/JBJS.I.01763.

24. Lyon T, Scheele W, Bhandari M, Koval KJ, Sanchez EG, Christensen J, et al. Efficacy and safety of recombinant human bone morphogenetic protein-2/calcium phosphate matrix for closed tibial diaphyseal fracture: A double-blind, randomized, controlled phase-II/III trial. J Bone Joint Surg Am 2013;95:2088-96. doi: 10.2106/JBJS.L.01545.

25. Burkus JK, Heim SE, Gornet MF, Zdeblick TA. Is INFUSE bone graft superior to autograft bone? An integrated analysis of clinical trials using the LT-CAGE lumbar tapered fusion device. J Spinal Disord Tech 2003;16:113-22. doi: 10.1097/00024720-200304000-00001.

26. Burkus JK, Gornet MF, Dickman CA, Zdeblick TA. Anterior lumbar interbody fusion using rhBMP-2 with tapered interbody cages. J Spinal Disord Tech 2002;15:337-49. doi: 10.1097/00024720-200210000-00001.

27. Haid RW Jr, Branch CL Jr, Alexander JT, Burkus JK. Posterior lumbar interbody fusion using recombinant human bone morphogenetic protein type 2 with cylindrical interbody cages. Spine J 2004;4:527-38. doi: 10.1016/j.spinee.2004.03.025.

28. Cho JH, Lee JH, Yeom JS, Chang BS, Yang JJ, Koo KH, 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. doi: 10.1016/j.spinee.2017.06.023.

29. Choi JY, Park HJ, Park SM, Kang CN, Song KS. Evaluation of the efficacy and safety of Escherichia coli-derived recombinant human bone morphogenetic protein-2 in transforaminal lumbar interbody fusion to treat degenerative spinal disease: A protocol of prospective, randomized controlled, assessor-blinded, open-label, multicenter trial. J Orthop Surg Res 2022;17:397. doi: 10.1186/s13018-022-03289-w.

30. Fiorellini JP, Howell TH, Cochran D, Malmquist J, Lilly LC, Spagnoli D, et al. Randomized study evaluating recombinant human bone morphogenetic protein-2 for extraction socket augmentation. J Periodontol 2005;76:605-13. doi: 10.1902/jop.2005.76.4.605.

31. Coomes AM, Mealey BL, Huynh-Ba G, Barboza-Arguello C, Moore WS, Cochran DL. Buccal bone formation after flapless extraction: A randomized, controlled clinical trial comparing recombinant human bone morphogenetic protein 2/absorbable collagen carrier and collagen sponge alone. J Periodontol 2014;85:525-35. doi: 10.1902/jop.2013.130207.

32. Wei L, Sun Y, Yu D, Pieterse H, Wismeijer D, Liu Y, et al. The clinical efficacy and safety of ErhBMP-2/BioCaP/β-TCP as a novel bone substitute using the tooth-extraction-socket-healing model: A proof-of-concept randomized controlled trial. J Clin Periodontol 2025;52:299-309. doi: 10.1111/jcpe.14084.

33. Jo DW, Cho YD, Seol YJ, Lee YM, Lee HJ, Kim YK. A randomized controlled clinical trial evaluating efficacy and adverse events of different types of recombinant human bone morphogenetic protein-2 delivery systems for alveolar ridge preservation. Clin Oral Implants Res 2019;30:396-409. doi: 10.1111/clr.13423.

34. Dickinson BP, Ashley RK, Wasson KL, O'Hara C, Gabbay J, Heller JB, et al. Reduced morbidity and improved healing with bone morphogenic protein-2 in older patients with alveolar cleft defects. Plast Reconstr Surg 2008;121:209-17. doi: 10.1097/01.prs.0000293870.64781.12.

35. Canan LW Jr, da Silva Freitas R, Alonso N, Tanikawa DY, Rocha DL, Coelho JC. Human bone morphogenetic protein-2 use for maxillary reconstruction in cleft lip and palate patients. J Craniofac Surg 2012;23:1627-33. doi: 10.1097/SCS.0b013e31825c75ba.

36. Alonso N, Tanikawa DY, Freitas Rda S, Canan L Jr, Ozawa TO, Rocha DL. Evaluation of maxillary alveolar reconstruction using a resorbable collagen sponge with recombinant human bone morphogenetic protein-2 in cleft lip and palate patients. Tissue Eng Part C Methods 2010;16:1183-9. doi: 10.1089/ten.TEC.2009.0824.

37. Schmidt AH. Autologous bone graft: Is it still the gold standard? Injury 2021;52 Suppl 2:S18-22. doi: 10.1016/j.injury.2021.01.043.

38. Chen Y, Xu C, Wu Y, Shi J, Chen R. Application of RhBMP-2 in percutaneous endoscopic posterior lumbar interbody fusion. BMC Surg 2024;24:376. doi: 10.1186/s12893-024-02674-y.

39. Uribe F, Vásquez B, Alister JP, Olate S. Comparison of rhBMP-2 in combination with different biomaterials for regeneration in rat calvaria critical-size defects. Biomed Res Int 2022;2022:6281641. doi: 10.1155/2022/6281641.

40. Lee J, Kang J, Seol J, Kim NS, Young HS. Effect of keratin-based biocomposite hydrogels as a RhBMP-2 carrier in calvarial bone defects mouse model. J Vet Clin 2022;39:302-10.

41. Fitzgerald A, McCool R, Carr E, Miller P, Reddish K, Lohr CC, et al. A systematic review of bone graft products used in lumbar interbody fusion procedures for degenerative disc disease. N Am Spine Soc J 2025;21:100579. doi: 10.1016/j.xnsj.2024.100579.

42. Alavi SA, Imanian M, Alkaabi S, Al-Sabri G, Forouzanfar T, Helder M. A systematic review and meta-analysis on the use of regenerative graft materials for socket preservation in randomized clinical trials. Oral Surg Oral Med Oral Pathol Oral Radiol 2024;138:702-18. doi: 10.1016/j.oooo.2024.07.003.

43. Konev VA, Labutin DV, Bozhkova SA. Experimental justification for clinical application of bone growth stimulators in traumatology and orthopaedics (A review). Sib Med Rev 2021;4:5-17.

44. Wijaya JH, Tjahyanto T, Alexi R, Purnomo AE, Rianto L, Arjuna YYE, Tobing et al. Application of rhBMP in spinal fusion surgery: Any correlation of cancer incidence? A systematic review and meta-analysis. Eur Spine J 2023;32:2020-8. doi: 10.1007/s00586-023-07730-4.

45. Shayeb MA, Elfadil S, Abutayyem H, Shqaidef A, Marrapodi MM, Cicciù M, et al. Bioactive surface modifications on dental implants: A systematic review and meta-analysis of osseointegration and longevity. Clin Oral Investig 2024;28:592. doi: 10.1007/s00784-024-05958-y.

46. Marshall KM, McLaren JS, Wojciechowski JP, Callens SJP, Echalier C, Kanczler JM, et al. Bioactive coatings on 3D printed scaffolds for bone regeneration: Use of Laponite® to deliver BMP-2 in an ovine femoral condyle defect model. Biomater Adv 2024;164:213959. doi: 10.1016/j.bioadv.2024.213959.

47. Zhao QT, Lai RF, Wang J, Li H. Experimental studies regarding injectable chitosan/β-TCP/rhBMP-2 composite in repairing rabbit mandible defects. CRTER 2009;13:10065-8.