Indonesian Journal of Global Health Research
Volume 8 Number 6, December 2026
e-ISSN 2715-1972; p-ISSN 2714-9749
http://jurnal.globalhealthsciencegroup.com/index.php/IJGHR

DOES BONE MARROW–BASED AUGMENTATION IMPROVE FUSION OUTCOMES IN LUMBAR SPINAL FUSION? A SYSTEMATIC REVIEW OF RANDOMIZED CONTROLLED TRIALS

Muhammad Ridho Akbar Eljatin[1], Teuku Nanta Aulia[2,3]*
[1]Department of Public Health, Faculty of Medicine, Universitas Syiah Kuala, Jl. Teungku Tanoh Abee, Kopelma Darussalam, Syiah Kuala, Banda Aceh, Aceh 24415, Indonesia
[2]Division of Orthopedics and Traumatology, Department of Surgery, Faculty of Medicine, Universitas Syiah Kuala, Jl. Tgk. Tanoh Abee, Kopelma Darussalam, Syiah Kuala, Banda Aceh, Aceh 24415, Indonesia
[3]Division of Orthopedics and Traumatology, Department of Surgery, Dr. Zainoel Abidin Hospital, Jl. Moh. Daud Beureuh No.108, Bandar Baru, Kuta Alam, Banda Aceh, Aceh 24415, Indonesia
*t.nantalbfk@usk.ac.id

ABSTRACT
Lumbar spinal fusion (LSF) is widely used for degenerative lumbar disorders, but pseudarthrosis remains a clinically relevant complication. Bone marrow–based augmentation has been proposed to enhance fusion while reducing reliance on conventional grafting. However, its comparative efficacy and safety remain uncertain. To evaluate the effects of bone marrow–based augmentation on radiographic fusion, clinical outcomes, perioperative outcomes, and safety following LSF. This systematic review followed PRISMA guidelines. PubMed/MEDLINE, Embase, Scopus, Web of Science, and Cochrane CENTRAL were searched without publication-year restrictions for randomized controlled trials (RCTs) evaluating bone marrow aspirate, bone marrow concentrate, or bone marrow-derived mesenchymal stromal cells in adults undergoing lumbar spinal fusion for degenerative lumbar disorders. Of 366 records identified, seven RCTs published between 2008 and 2023, involving 367 participants, met the eligibility criteria. Data on radiographic fusion, clinical and functional outcomes, perioperative outcomes, and safety were extracted and synthesized narratively because of substantial clinical and methodological heterogeneity. Risk of bias was assessed using the Cochrane Risk of Bias 2 tool. Seven RCTs involving 367 participants demonstrated heterogeneous effects on radiographic fusion. Higher fusion rates were observed with selected constructs, including bone marrow-derived mesenchymal stromal cells plus allograft, bone marrow concentrate plus allograft, and bone marrow aspirate plus local bone and hydroxyapatite, whereas other studies showed no significant benefit or lower fusion rates with specific graft combinations. Clinical outcomes generally improved in both groups without consistent between-group differences. Perioperative outcomes and reported safety events were broadly comparable, although safety reporting was inconsistent. Four studies had some concerns, and three had a high overall risk of bias. Bone marrow–based augmentation showed variable effects on fusion, with benefits observed only in selected graft constructs. Current evidence does not establish consistent superiority over conventional grafting for LSF.

Keywords: bone marrow aspirate; bone marrow concentrate; mesenchymal stromal cell; lumbar fusion; pseudarthrosis; spinal arthrodesis

How to cite (in APA style)
Eljatin, M. R. A., & Aulia, T. N. (2026). Does Bone Marrow–Based Augmentation Improve Fusion Outcomes in Lumbar Spinal Fusion? A Systematic Review of Randomized Controlled Trials. *Indonesian Journal of Global Health Research*, *8*(6), 559–572. https://doi.org/10.37287/ijghr.v8i6.2729.

INTRODUCTION

Lumbar spinal fusion is an established surgical treatment for symptomatic degenerative lumbar disorders, including degenerative disc disease, spinal stenosis, and spondylolisthesis, with successful arthrodesis being an important component of treatment success (García de Frutos et al., 2020; Johnson, 2014). Pseudarthrosis remains a clinically relevant complication because non-union may contribute to persistent pain, mechanical instability, implant failure, and revision surgery (Chotivichit et al., 2016; Niu et al., 2009). Autologous iliac crest bone graft (ICBG) has traditionally been regarded as the reference standard because it provides osteogenic cells, osteoinductive factors, and an osteoconductive scaffold (Pokharel et al., 2022). However, iliac crest harvesting is associated with donor-site pain and other morbidity, while the quantity and biological

quality of available autograft may be limited (Hart et al., 2014; Johnson, 2014). These limitations have stimulated the development of alternative graft materials and biological augmentation strategies designed to maintain fusion efficacy while reducing donor-site morbidity.

Bone marrow-based augmentation has attracted particular interest because it contains osteogenic progenitor cells, including mesenchymal stromal cells (MSCs), along with biologically active factors that may contribute to bone formation and graft incorporation (Ambrosio et al., 2025; Areesak Chotivichit et al., 2016; Hart et al., 2014; Johnson, 2014). Bone marrow can be used as aspirate or concentrated marrow and combined with local bone, allograft, or osteoconductive scaffolds, providing a biological source of progenitor cells while the graft material serves as a structural substrate for new bone formation (Hart et al., 2014; Niu et al., 2009). Clinical evidence, however, has produced inconsistent findings. Johnson et al. reported that cancellous allograft combined with bone marrow concentrate achieved fusion equivalent to autologous iliac crest graft, whereas Hart et al. demonstrated significantly greater fusion with bone marrow concentrate plus allograft than with allograft alone (Hart et al., 2014; Johnson, 2014). Conversely, Chotivichit et al. found that adding bone marrow concentrate to local bone did not improve fusion and was associated with a high proportion of incomplete bridging on CT assessment (Chotivichit et al., 2016).

Other clinical studies further show that outcomes may vary by the cellular product and scaffold used. Niu et al. reported that bone marrow aspirate combined with laminectomy bone produced fusion comparable with ICBG, whereas calcium sulfate pellets combined with marrow aspirate produced substantially lower fusion, suggesting that scaffold characteristics may influence the biological effect of marrow augmentation (Niu et al., 2009). More advanced approaches have also been investigated, including expanded autologous BM-MSCs combined with allogeneic bone and cellular bone allografts, with clinical studies reporting favorable radiographic and clinical outcomes (García de Frutos et al., 2020; Park et al., 2023). Nevertheless, differences in marrow preparation, graft composition, surgical technique, fusion level, follow-up duration, and radiographic criteria make the overall efficacy of bone marrow-based augmentation uncertain .

Therefore, the available evidence suggests that the effect of bone marrow augmentation may depend not only on the presence of marrow-derived cells but also on the graft substrate and clinical context in which it is applied. Across studies, results have varied rather than converged on a single effect, indicating the need for a focused synthesis to determine whether bone marrow-based augmentation provides a consistent benefit over conventional grafting and whether such benefit extends beyond radiographic fusion to clinical, perioperative, and safety outcomes. This systematic review aimed to evaluate the effects of bone marrow-based augmentation on radiographic fusion, clinical outcomes, perioperative outcomes, and safety in patients undergoing lumbar spinal fusion based exclusively on randomized controlled trials.

METHOD

Study design and reporting

This systematic review evaluated the effects of bone marrow–based augmentation on radiographic fusion and clinical outcomes following lumbar spinal fusion. The review was reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement. We conducted the literature search, study selection, and data extraction between May and August 2025.

Literature search

A systematic search was performed in five electronic databases: PubMed/MEDLINE, Embase, Scopus, Web of Science, and the Cochrane Central Register of Controlled Trials (CENTRAL). The search strategy was developed to capture studies evaluating bone marrow–based biological augmentation in lumbar spinal fusion. Three main concepts were combined using Boolean

operators: lumbar spinal fusion, bone marrow–based interventions, and stem/progenitor cell-based augmentation.

For PubMed/MEDLINE, Medical Subject Headings (MeSH) and free-text terms were used. We adapted equivalent free-text terms and database-specific indexing terms for Embase, Scopus, Web of Science, and CENTRAL. Search terms included combinations of "lumbar spinal fusion," "lumbar fusion," "spinal fusion," "posterolateral fusion," "bone marrow aspirate," "bone marrow concentrate," "bone marrow aspirate concentrate," "bone marrow-derived mesenchymal stem cells," "mesenchymal stem cells," "stem cell," and "autologous." Boolean operators AND and OR were used to combine search terms.

No restriction on publication year was applied. The search was limited to English-language studies. The seven RCTs included in the final synthesis were published between 2008 and 2023. Table 1 provides the complete search strategy for each database, including the search terms and database-specific syntax. We also screened the reference lists of eligible studies manually to identify potentially relevant articles that may not have been retrieved through the electronic database search.

Table 1.
Database keywords

Database	Search Terms
PubMed/MEDLINE	("lumbar spinal fusion" OR "lumbar fusion" OR "spinal fusion") AND ("bone marrow aspirate" OR "bone marrow concentrate" OR "bone marrow aspirate concentrate" OR "mesenchymal stem cells") AND ("autologous" OR "bone marrow")
Embase	("lumbar spinal fusion" OR "lumbar fusion" OR "spinal fusion") AND ("bone marrow aspirate" OR "bone marrow concentrate" OR "bone marrow aspirate concentrate" OR "mesenchymal stem cells") AND ("autologous" OR "bone marrow")
Scopus	("lumbar spinal fusion" OR "lumbar fusion" OR "spinal fusion") AND ("bone marrow aspirate" OR "bone marrow concentrate" OR "bone marrow aspirate concentrate" OR "mesenchymal stem cells") AND ("autologous" OR "bone marrow")
Web of Science	("lumbar spinal fusion" OR "lumbar fusion" OR "spinal fusion") AND ("bone marrow aspirate" OR "bone marrow concentrate" OR "bone marrow aspirate concentrate" OR "mesenchymal stem cells") AND ("autologous" OR "bone marrow")
Cochrane CENTRAL	("lumbar spinal fusion" OR "lumbar fusion" OR "spinal fusion") AND ("bone marrow aspirate" OR "bone marrow concentrate" OR "bone marrow aspirate concentrate" OR "mesenchymal stem cells") AND ("autologous" OR "bone marrow")

Eligibility criteria

Eligibility criteria were established before study selection and were based on the population, intervention, comparator, outcomes, and study design of the review. Studies were eligible if they met the following criteria: (1) human clinical studies involving patients undergoing lumbar spinal fusion for degenerative lumbar spinal disorders; (2) randomized controlled trials evaluating a bone marrow–based augmentation strategy, including bone marrow aspirate (BMA), bone marrow concentrate (BMC), bone marrow aspirate concentrate (BMAC), or bone marrow-derived mesenchymal stem cells (BM-MSCs); (3) comparison with a conventional grafting strategy or corresponding graft material without bone marrow augmentation; and (4) reporting of radiographic fusion outcomes as the primary outcome. Clinical, functional, perioperative, and safety outcomes were considered secondary outcomes when available.

Eligible surgical procedures included lumbar fusion procedures performed for degenerative conditions, including degenerative disc disease, degenerative spondylolisthesis, lumbar spinal stenosis, spondylolytic spondylolisthesis, and segmental instability. Studies were eligible regardless of the number of lumbar levels treated. The intervention was defined as a bone marrow–based biological augmentation strategy, including direct BMA, concentrated bone marrow products, or ex vivo-expanded autologous BM-MSCs, used either alone or in combination with local bone, autograft, allograft, hydroxyapatite, calcium sulfate, or other grafting materials.

Studies were excluded if they (1) were non-randomized studies; (2) were conducted in animals, cadaveric models, or laboratory settings; (3) evaluated non-lumbar spinal fusion procedures; (4) investigated allogeneic or xenogeneic cell products without an eligible autologous bone marrow intervention; or (5) were reviews, systematic reviews, meta-analyses, case reports, case series, editorials, technical notes, conference abstracts, letters, or publications without primary comparative clinical data. Articles published in languages other than English were also excluded.

Study selection

All records identified through the electronic database searches were compiled and screened for duplicate records. After removing duplicates, titles and abstracts were screened against the predefined eligibility criteria. Potentially relevant records were subsequently assessed by full-text review. Full-text articles that did not meet the eligibility criteria were excluded, and the reasons for exclusion were recorded. The study-selection process was summarized using a PRISMA 2020 flow diagram, including the numbers of records identified, duplicates removed, records screened, reports assessed for eligibility, reports that could not be retrieved, full-text reports excluded, and studies included in the final synthesis.

Table 2.
PICOS criteria for study selection

Population	Adult patients undergoing lumbar spinal fusion for degenerative lumbar disorders, including degenerative disc disease, degenerative spondylolisthesis, lumbar spinal stenosis, spondylolytic spondylolisthesis, and segmental instability.
Intervention	Bone marrow–based augmentation, including bone marrow aspirate (BMA), bone marrow concentrate (BMC), or ex vivo-expanded bone marrow-derived mesenchymal stromal cells (BM-MSCs), used alone or combined with local bone, allograft, autograft, hydroxyapatite, calcium sulfate, or other graft materials.
Comparison	Conventional grafting strategies without bone marrow augmentation, including iliac crest bone graft (ICBG), local bone graft, allograft, or corresponding graft material alone.
Outcome	Primary outcome: Radiographic fusion assessed by plain radiography, computed tomography (CT), or three-dimensional CT. Secondary outcomes: Clinical and functional outcomes, including Oswestry Disability Index (ODI), Visual Analog Scale (VAS), and 36-Item Short Form Health Survey (SF-36); perioperative outcomes, including blood loss, operative time, and length of hospital stay; and safety outcomes, including adverse events, surgical-site infection, donor-site morbidity, neurological complications, and reoperation.
Study Design	Randomized controlled trials (RCTs), including prospective parallel-group and within-patient randomized designs.

Data extraction

Data were extracted using a predefined data-extraction form. The following information was collected from each eligible study: first author and publication year, country, study design, sample size, patient population and underlying diagnosis, spinal level and number of levels treated, surgical procedure, bone marrow–based intervention, graft composite, comparator intervention, fusion assessment method, follow-up duration, radiographic fusion outcomes, clinical and functional outcomes, perioperative outcomes, and adverse events or other safety outcomes. For fusion outcomes, the original definitions and assessment methods reported by each study were retained. Radiographic assessment included plain radiography, computed tomography (CT), and three-dimensional CT. Data were extracted separately for different follow-up time points when multiple assessments were reported.

Outcomes

The primary outcome was radiographic fusion following lumbar spinal fusion. Fusion was defined according to the criteria used in each original study, including complete fusion, solid fusion, complete bony bridging, continuous intertransverse bony bridging, or study-specific radiographic fusion scores. Secondary outcomes included clinical and functional measures, particularly the

Oswestry Disability Index (ODI), Visual Analog Scale (VAS), and 36-Item Short Form Health Survey (SF-36). Perioperative outcomes included operative time, intraoperative blood loss, and length of hospital stay. Safety outcomes included surgical-site infection, treatment-emergent adverse events (TEAEs), serious TEAEs, donor-site morbidity, aspiration-site complications, neurological complications, reoperation, and other adverse events reported by the individual studies.

Risk of bias assessment

We assessed risk of bias using the Cochrane Risk of Bias 2 (RoB 2) tool for randomized trials. The assessment covered five domains: bias arising from the randomization process (D1), bias due to deviations from intended interventions (D2), bias due to missing outcome data (D3), bias in outcome measurement (D4), and bias in selection of the reported result (D5). Each domain was categorized as low risk, some concerns, or high risk of bias, according to the RoB 2 framework. The overall risk of bias for each study was determined based on the domain-level assessments. Table 4 presents the complete assessment.

Data synthesis

Study characteristics and outcome data were initially synthesized descriptively. Given the substantial clinical and methodological heterogeneity among the included trials—including differences in bone marrow product, graft composite, surgical procedure, fusion assessment modality, follow-up duration, and comparator intervention—the primary synthesis was narrative rather than quantitative. Fusion findings were compared according to the type of bone marrow–based intervention and the corresponding control intervention, while preserving the outcome definitions and follow-up intervals reported by the original investigators. Clinical, perioperative, and safety outcomes were similarly summarized according to the available data.

Assessment of heterogeneity

Heterogeneity was assessed qualitatively by considering differences in patient populations, underlying lumbar pathology, surgical technique, number of fused levels, bone marrow product, graft material, comparator, radiographic assessment method, and follow-up duration. Particular attention was given to differences between BMA, BMC, and ex vivo-expanded BM-MSC interventions and to differences in the graft materials with which these products were combined. These methodological differences were considered when interpreting the variability in fusion outcomes across the included trials.

Reporting of missing or unavailable information

When an outcome was not reported by an included study, it was recorded as not reported (NR) rather than inferred or estimated. This was particularly relevant to Niu et al. (2009), which explicitly evaluated and recorded fusion status without reporting clinical or adverse-event outcomes.

RESULT

Study selection and characteristics

The electronic search of five databases—PubMed/MEDLINE, Embase, Scopus, Web of Science, and Cochrane CENTRAL—identified 366 records. After removal of nine duplicates, 357 records underwent title and abstract screening, and 329 were excluded because they were not relevant to the research question or did not evaluate bone marrow-based interventions for lumbar spinal fusion. Twenty-eight reports were assessed for full-text eligibility, and two could not be retrieved. The remaining 26 reports were assessed for eligibility, and 19 were excluded because of ineligible study design, unrelated research scope, or insufficient methodological information. Seven RCTs met the predefined eligibility criteria and were included in the qualitative synthesis (Figure 1).

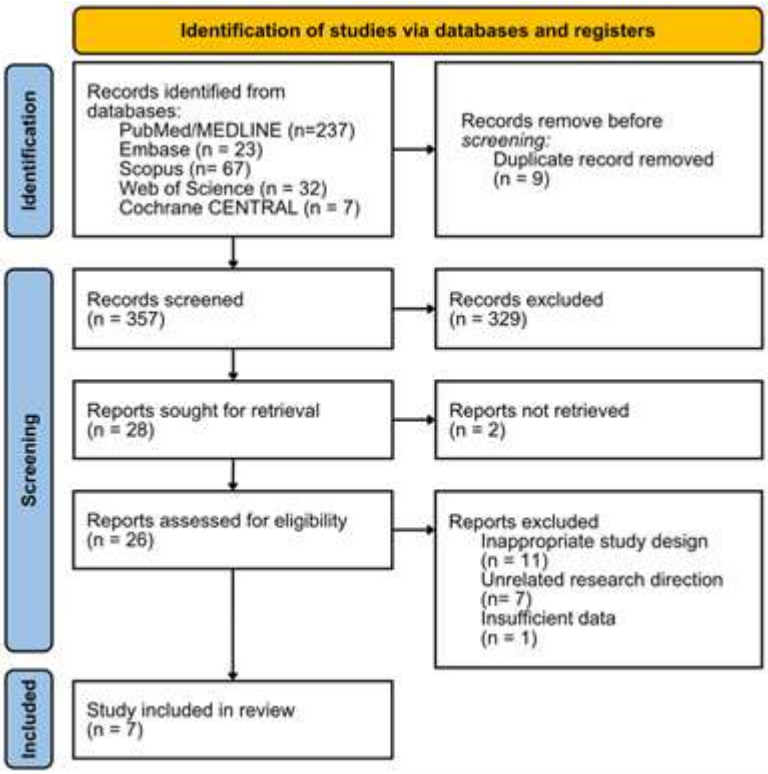

Figure 1. PRISMA flowchart of article search and selection

The seven studies were published between 2008 and 2023 and conducted in Thailand, Spain, the Czech Republic, the USA, and Taiwan (Table 1). Sample sizes ranged from 12 to 99 participants. Most studies used a prospective parallel-group RCT design, whereas Johnson (2014) and Niu et al. (2009) used within-patient comparisons. Interventions included BMA, BMC, and ex vivo-expanded BM-MSCs combined with LBG, allograft, HA, or CS, and control groups received ICBG, allograft, or the corresponding graft material without bone marrow augmentation. Follow-up ranged from 6 to 24 months.

In the 2023 study by Sumethvanich and Paiboon, 99 patients entered the study, with 50 allocated to the BMA group and 49 to the control group; after losses to follow-up and incomplete CT assessment, 95 patients remained for analysis (48 vs 47). García de Frutos et al. (2020) randomized 73 adults with grade I–II L4–L5 DS and/or DDD undergoing TLIF to receive ex vivo-expanded BM-MSCs with cancellous allograft or ICBG. Chotivichit et al. (2016) included 12 patients with single-level L4–L5 DS undergoing instrumented PLF and compared BMC plus LBG with LBG alone. Hart et al. (2014) enrolled 80 patients with lumbar degenerative disease undergoing instrumented lumbar or lumbosacral PLF and compared BMC plus spongious allograft with allograft alone. Johnson (2014) enrolled 25 patients with DDD, DS, or LS undergoing one- to three-level lumbar fusion and used a within-patient design comparing BMC plus allograft with ICBG. Niu et al. (2009) included 43 patients undergoing one-level instrumented fusion, comparing ALB plus BMA or CS plus BMA with ICBG on the contralateral side. Pantarote and Prasopchoke (2008) included 40 patients undergoing two-level instrumented PLF and compared BMA plus local bone with ICBG.

Fusion outcomes

Fusion outcomes are summarised in Table 2. Findings varied across studies, with effects differing by bone marrow product, graft composite, surgical procedure, and follow-up period. Sumethvanich and Paiboon (2023) reported a significantly higher CT-defined fusion rate with BMA plus LBG and HA than with LBG and HA alone at 6 months (39.6% vs 12.8%, p=0.005). At 12 months, the difference was no longer statistically significant (75.0% vs 55.3%, p=0.054). In the subgroup

undergoing 2–3-level fusion, however, the 12-month fusion rate remained significantly higher with BMA (85.2% vs 58.1%, p=0.042). García de Frutos et al. (2020) reported significantly higher CT-defined posterior fusion with BM-MSCs plus allograft at 6 months (82.35% vs 45.71%, p=0.0001). At 12 months, the difference was not statistically significant (76.47% vs 65.71%, p=0.0731). Complete response, defined by the study as the presence of both posterior intertransverse and anterior interbody fusion, was also higher with BM-MSCs at 6 months (70.59% vs 40.00%, p=0.0038) and 12 months (70.59% vs 51.43%, p=0.0229). The published abstract reported 92.3% versus 45.7% for CT-based posterior fusion at 6 months, whereas the detailed results and Table 5 reported 82.35% versus 45.71%; the latter values are retained in Table 2.

Chotivichit et al. (2016) found a significantly lower 6-month complete PLF bridging rate with BMC plus LBG than with LBG alone (58.3% vs 100%, p=0.01). In contrast, Hart et al. (2014) reported higher fusion with BMC plus allograft than with allograft alone. X-ray-defined complete fusion was 15% vs 0% at 12 months (p=0.041) and 35% vs 10% at 24 months (p=0.011), while CT at 24 months showed continuous bridging bone in 80% vs 40% (p=0.003). Johnson (2014) found no significant difference between BMC plus allograft and ICBG in lateral gutters or facet joints. The interbody fusion score was also not significantly different (0.7±0.2 vs 1.0±0.2, p>0.05). Niu et al. (2009) reported comparable fusion with ALB plus BMA and ICBG (85.7% vs 90.5%, p=0.56). For CS plus BMA, the detailed Results reported significantly lower fusion than ICBG (40.5% vs 90.9%, p=0.0016). The abstract reported 45.5% rather than 40.5% for the CS+BMA group; this discrepancy was retained rather than recalculated. Pantarote and Prasopchoke (2008) reported no significant difference in 12-month fusion between BMA plus local bone and ICBG (85% vs 90%, p>0.05).

Table 4.
Fusion outcomes following bone marrow-based interventions

Study	Assessment method	Follow-up	Intervention	Control	Fusion outcome
Sumethvanich & Paiboon, 2023	CT	6 months	BMA + LBG + HA	LBG + HA	39.6% vs 12.8%, p=0.005
		12 months			75.0% vs 55.3%, p=0.054
		12 months, 2–3 levels			85.2% vs 58.1%, p=0.042
García de Frutos et al., 2020	CT	6 months	BM-MSCs + allograft	ICBG	82.35% vs 45.71%, p=0.0001
		12 months			76.47% vs 65.71%, p=0.0731
		6 months, complete response			70.59% vs 40.00%, p=0.0038
		12 months, complete response			70.59% vs 51.43%, p=0.0229
Chotivichit et al., 2016	3D-CT	6 months	BMC + LBG	LBG	58.3% vs 100%, p=0.01
Hart et al., 2014	X-ray	12 months	BMC + allograft	Allograft	15% vs 0%, p=0.041
	CT	24 months			80% vs 40%, p=0.003
	X-ray	24 months			35% vs 10%, p=0.011
Johnson, 2014	CT	12 months	BMC + allograft	ICBG	No significant difference; fusion score 0.7±0.2 vs 1.0±0.2, p>0.05
Niu et al., 2009	Radiography/CT	≥12 months	ALB + BMA	ICBG	85.7% vs 90.5%, p=0.56
		≥12 months	CS + BMA	ICBG	40.5% vs 90.9%, p=0.0016
Pantarote & Prasopchoke, 2008	X-ray	12 months	BMA + local bone	ICBG	85% vs 90%, p>0.05

ALB, autogenous laminectomy bone; BMA, bone marrow aspirate; BMC, bone marrow concentrate; BM-MSCs, bone marrow-derived mesenchymal stem cells; CS, calcium sulfate; CT, computed tomography; HA, hydroxyapatite; ICBG, iliac crest bone graft; LBG, local bone graft. Overall, the fusion effect was heterogeneous across the seven studies. Bone marrow-based augmentation was associated with higher fusion in the 2023 Sumethvanich study and the 2020 García de Frutos study, lower fusion in Chotivichit et al. and with CS plus BMA in Niu et al., no significant difference in Johnson and Pantarote, and higher fusion with BMC plus allograft in Hart et al.

Clinical, perioperative, and safety outcomes

Clinical, perioperative, and safety outcomes are summarised in Table 3. García de Frutos et al. (2020) provided the most extensive patient-reported outcome data. Across outcomes, both groups generally improved from baseline, with no significant between-group difference in change. SF-36 MCS improved significantly from baseline in the BM-MSC group but not consistently in the control group. However, the between-group differences in change were not statistically significant at 3, 6, or 12 months. Perioperative outcomes in Sumethvanich and Paiboon (2023) were comparable between groups. Blood loss, operative time, length of hospital stay, and superficial and deep SSI did not differ significantly. Hart et al. (2014) also reported similar hospital stays between groups (11.8 vs 12.0 days, p=0.72).

Table 3.
Clinical, perioperative, and safety outcomes

Study	Outcome	Follow-up	Intervention	Control	Between-group result
Sumethvanich & Paiboon, 2023	Intraoperative blood loss, median (IQR)	Perioperative	BMA + LBG + HA	LBG + HA	375 (225–625) vs 400 (200–700) mL, p=0.890
	Operative time, mean (SD)	Perioperative			212 (54) vs 210 (46) min, p=0.833
	Length of hospital stay, mean (SD)	Postoperative			8.4 (2.8) vs 7.7 (3.3) days, p=0.305
	Superficial SSI	Postoperative			6.3% vs 2.1%, p=0.617
	Deep SSI	Postoperative			2.1% vs 2.1%, p=1.000
García de Frutos et al., 2020	Lumbar VAS	3, 6, and 12 months	BM-MSCs + allograft	ICBG	Significant improvement in both groups; no significant between-group difference
	Sciatic VAS	3, 6, and 12 months			Significant improvement in both groups; no significant between-group difference
	ODI	3, 6, and 12 months			Significant improvement in both groups; no significant between-group difference
	SF-36 PCS	3, 6, and 12 months			Significant improvement in both groups; no significant between-group difference
	SF-36 MCS	3, 6, and 12 months			No significant between-group difference in change from baseline
	Any TEAE	Study period			88.24% vs 97.14%
	Serious TEAE	Study period			14.71% vs 8.57%
	TEAE leading to death	Study period			0% vs 2.86%
	TEAE related to experimental treatment	Study period			0% vs 0%
Chotivichit et al., 2016	ODI	Baseline, 3, and 6 months	BMC + LBG	LBG	No significant between-group difference
	Fusion-mass volume	1 week and 6 months			No significant between-group difference

	BM aspiration-site inflammation	Postoperative			1 patient vs 0 patients
Hart et al., 2014	Hospital stay, mean	Postoperative	BMC + allograft	Allograft	11.8 vs 12.0 days, p=0.72
	Hematoma requiring revision/drainage	First postoperative week			2 vs 2 patients
	Further surgery for any reason	24 months			0 vs 0 patients
	Other AEs	24 months			NR
Johnson, 2014	Fusion score, lateral gutters/facet joints	12 months	BMC + allograft	ICBG	No significant difference, p>0.05
	Interbody fusion score, mean (SD)	12 months			0.7 ± 0.2 vs 1.0 ± 0.2, p>0.05
	Intraoperative/postoperative complications related to BMA	12 months			NR
	Infection	12 months			NR
	Neurological deficit	12 months			NR
	Return to operating room	12 months			NR
Niu et al., 2009	Clinical outcome	—	ALB + BMA / CS + BMA	ICBG	NR
	Adverse events	—			NR
Pantarote & Prasopchoke, 2008	Donor-site pain	12 months	BMA + local bone	ICBG	0 vs 1 patient, p>0.05
	Iliac crest fracture	12 months			0 vs 2 patients, p>0.05
	Overall donor-site morbidity	12 months			0 vs 3 patients, p>0.05

AE, adverse event; ALB, autogenous laminectomy bone; BMA, bone marrow aspirate; BMC, bone marrow concentrate; BM-MSCs, bone marrow-derived mesenchymal stem cells; CS, calcium sulfate; HA, hydroxyapatite; ICBG, iliac crest bone graft; IQR, interquartile range; LBG, local bone graft; NR, not reported; ODI, Oswestry Disability Index; PCS, physical component summary; MCS, mental component summary; SD, standard deviation; SSI, surgical site infection; TEAE, treatment-emergent adverse event; VAS, visual analog scale.

Safety outcomes were generally comparable where reported. Across studies, TEAEs, serious TEAEs, and TEAEs leading to death were not clearly different between groups, and no TEAE was considered related to the experimental treatment. Chotivichit et al. reported local inflammation at the BM aspiration site in one patient in the BMC group and none in the control group. Hart et al. reported two hematomas requiring revision or drainage in each group and no further surgery during follow-up. Johnson et al. reported no significant difference in radiographic fusion scores and no neurological deficit or return to the operating room. Pantarote and Prasopchoke reported no significant difference in donor-site pain, iliac crest fracture, or overall donor-site morbidity. Clinical outcomes and adverse events were not reported by Niu et al., because the study explicitly evaluated and recorded fusion status only.

Risk of bias

Risk of bias was assessed using the Cochrane RoB 2 tool across five domains: bias arising from the randomization process, deviations from intended interventions, missing outcome data, measurement of the outcome, and selection of the reported result (Table 4). Four studies were judged to have some concerns, whereas three studies were judged to have a high overall risk of bias. Sumethvanich and Paiboon (2023), García de Frutos et al. (2020), and Hart et al. (2014) had low risk in most domains but some concerns regarding deviations from intended interventions. Pantarote and Prasopchoke (2008) had some concerns regarding randomization, deviations from intended

interventions, and selection of the reported result. Chotivichit et al. (2016) was judged to have some concerns regarding randomization and deviations from intended interventions and high risk regarding outcome measurement. Johnson (2014) and Niu et al. (2009) were both judged to have high risk arising from the randomization process, resulting in a high overall risk of bias. The complete domain-level assessment is presented in Table 4.

Table 4.
Risk of bias assessment using the Cochrane RoB 2 tool

Study	D1	D2	D3	D4	D5	Overall risk of bias
Sumethvanich and Paiboon (2023)	Low	Some concerns	Low	Low	Low	Some concerns
García de Frutos et al. (2020)	Low	Some concerns	Low	Low	Low	Some concerns
Chotivichit et al. (2016)	Some concerns	Some concerns	Low	High	Low	High
Hart et al. (2014)	Low	Some concerns	Low	Low	Low	Some concerns
Johnson (2014)	High	Low	Low	Low	Low	High
Niu et al. (2009)	High	Low	Low	Low	Low	High
Pantarote and Prasopchoke (2008)	Some concerns	Some concerns	Low	Low	Some concerns	Some concerns

Domains: D1, bias arising from the randomisation process; D2, bias due to deviations from intended interventions; D3, bias due to missing outcome data; D4, bias in measurement of the outcome; D5, bias in selection of the reported result; RoB 2, revised Cochrane risk-of-bias tool for randomised trials.

Risk-of-bias categories: Low, low risk of bias; Some concerns, some concerns regarding risk of bias; High, high risk of bias.

Summary of findings

Across seven RCTs, bone marrow-based interventions showed heterogeneous effects on radiographic fusion. Higher fusion rates were reported at selected time points in the 2023 and 2020 studies, whereas BMC- or BMA-based strategies produced lower fusion rates in some comparisons. Other studies showed no significant difference from conventional grafting. Clinical outcomes were less frequently reported but generally improved in both intervention and control groups. In the most extensively reported trial, pain, disability, and physical health measures improved in both groups without a significant between-group difference in change. Perioperative outcomes and reported safety events were also broadly comparable between groups, although reporting was inconsistent across studies.

Overall, the included trials do not provide consistent evidence that bone marrow-based augmentation is superior to conventional grafting for lumbar spinal fusion. The variability in fusion findings across studies should be interpreted in the context of differences in intervention type, graft composite, surgical technique, follow-up duration, outcome assessment, and the methodological limitations identified in the RoB 2 assessment.

Table 3.
Characteristics and treatment protocols of included studies (n= 7)

Study	Country	Study Design (Sample Size)	Sample	Intervention Graft Composite	Control Graft Composite	Follow-up
Sumethvanich & Paiboon, 2023	Thailand	Prospective RCT (N = 99)	DLSS undergoing decompression laminectomy + instrumented PLF (2–6 levels)	BMA (15 mL, vertebral body) + LBG + HA	LBG + HA	12 months
García de Frutos et al.,	Spain	Prospective RCT (N =	Grade I–II L4–L5 DS and/or DDD undergoing TLIF	Ex vivo–expanded BM-	ICBG	12 months

Study	Country	Study Design (Sample Size)	Sample	Intervention Graft Composite	Control Graft Composite	Follow-up
2020		73)		MSCs + cancellous allograft		
Chotivichit et al., 2016	Thailand	Prospective RCT (N = 12)	Single-level L4–L5 DS undergoing decompressive laminectomy + instrumented PLF	BMC + LBG	LBG	6 months
Hart et al., 2014	Czech Republic	Prospective RCT (N = 80)	Lumbar degenerative disease undergoing instrumented lumbar/lumbosacral PLF (1–3 levels)	BMC (10 mL) + spongious allograft chips (90 g)	Spongious allograft chips (90 g)	24 months
Johnson, 2014	USA	Prospective within-patient RCT (N = 25)	DDD, DS, or LS undergoing lumbar fusion (1–3 levels)	BMC (16 mL) + cancellous allograft (30 mL) + thrombin + $CaCl_2$	ICBG	12 months
Niu et al., 2009	Taiwan	Prospective within-patient RCT (N = 43)	LSS with low-grade DS/SS (Grade I–II) and/or segmental instability undergoing 1-level instrumented PLF	Arm 1: ALB + BMA; Arm 2: CS + BMA	ICBG	24 months
Pantarote & Prasopchoke, 2008	Thailand	Prospective RCT (N = 40)	Spondylolisthesis with spinal stenosis undergoing 2-level instrumented PLF	BMA (15 mL) + LBG (20–25 cm^3)	ICBG	12 months

ALB, autogenous laminectomy bone; BMA, bone marrow aspirate; BMC, bone marrow concentrate; BM-MSCs, bone marrow–derived mesenchymal stromal cells; $CaCl_2$, calcium chloride; CS, calcium sulfate; DDD, degenerative disc disease; DLSS, degenerative lumbar spinal stenosis; DS, degenerative spondylolisthesis; HA, hydroxyapatite; ICBG, iliac crest bone graft; LS, lumbar stenosis; LBG, local bone graft; PLF, posterolateral lumbar fusion; RCT, randomized controlled trial; SS, spondylolytic spondylolisthesis; TLIF, transforaminal lumbar interbody fusion.

DISCUSSION

This systematic review of seven randomized controlled trials found inconsistent effects of bone marrow–based augmentation on radiographic fusion in lumbar spinal fusion. Across studies, selected fusion outcomes improved with expanded BM-MSCs combined with allograft (García de Frutos et al., 2020), BMC combined with allograft (Hart et al., 2014), and BMA added to local bone graft and hydroxyapatite (Sayun Sumethvanich & Anuwat Paiboon, 2023). However, BMC augmentation of local bone graft did not improve fusion in the small trial by Chotivichit et al. (2016), while Johnson (2014) and Pantarote and Prasopchoke (2008) reported no significant difference compared with ICBG. Niu et al. (2009) further showed a construct-dependent effect, with laminectomy bone plus BMA achieving fusion comparable to ICBG, whereas calcium sulfate plus BMA resulted in significantly lower fusion (Chotivichit et al., 2016; Johnson, 2014; Niu et al., 2009). Together, these findings indicate that bone marrow–based augmentation shows selective rather than universal fusion benefit.

The inconsistent findings likely reflect the substantial heterogeneity of the interventions evaluated. The included trials investigated unprocessed BMA, point-of-care BMC, and culture-expanded BM-MSCs, with substantial differences in marrow source, processing, cellular composition, and graft combination (García de Frutos et al., 2020; Hart et al., 2014; Johnson, 2014). The surgical context also varied from instrumented PLF to TLIF and from single-level to multilevel fusion, while radiographic assessment and follow-up ranged from 6 to 24 months. These differences make “bone marrow–based augmentation” a broad treatment category rather than a single standardized intervention. Gordon et al. (2024) also emphasized this issue, identifying considerable variation in cell preparation and grafting techniques across the literature (Gordon et al., 2024). Similarly, Salamanna et al. (2024) reported substantial variability in the use of autologous cellular approaches and graft substitutes for spinal fusion (Salamanna et al., 2024). Therefore, the observed heterogeneity in our review should be interpreted as an important characteristic of the intervention

itself rather than simply as inconsistency between individual trials.

The included trials suggest that the graft or scaffold accompanying bone marrow may be an important determinant of the observed effect. Across studies, Niu et al. (2009) found that laminectomy bone combined with BMA achieved 85.7% fusion compared with 90.5% with ICBG, whereas calcium sulfate combined with BMA achieved only 40.5% compared with 90.9% for ICBG. The authors emphasized that an appropriate osteoconductive scaffold is required to support osteogenic cells and osteoinductive factors (Niu et al., 2009). This construct-dependent pattern was also evident in Hart et al. (2014), in which adding BMC to cancellous allograft increased fusion from 0% to 15% at 12 months and from 10% to 35% at 24 months on radiographs, while CT at 24 months showed bridging bone in 80% versus 40% (Hart et al., 2014). In Sumethvanich and Paiboon (2023), BMA was used specifically with local bone graft and hydroxyapatite and produced higher fusion at 6 months (Sayun Sumethvanich & Anuwat Paiboon, 2023). Overall, these findings suggest that marrow-derived cellular augmentation may be better viewed as a component of a graft construct rather than an independent determinant of fusion. This interpretation is consistent with recent reviews emphasizing the variability in biological activity and grafting strategies among osteobiologics (Ambrosio et al., 2025; Lambrechts et al., 2024).

A key finding of this review is that improved radiographic fusion did not consistently translate into better clinical outcomes. García de Frutos et al. (2020) reported higher CT-based posterior fusion with MSCs plus allograft at 6 months (82.35% vs 45.71%), although the difference was no longer statistically significant at 12 months (76.47% vs 65.71%). Nevertheless, lumbar pain, sciatic pain, and ODI improved in both groups, with no significant between-group differences (García de Frutos et al., 2020). Chotivichit et al. (2016) similarly found no significant between-group difference in ODI despite lower radiographic fusion in the BMC group (Chotivichit et al., 2016). Johnson (2014) also reported equivalent radiographic fusion between BMC plus allograft and ICBG (Johnson, 2014). Gordon et al. (2024) support this distinction, finding no difference in functional outcome scores between cell-based grafts and ICBG despite variable fusion results (Gordon et al., 2024). More broadly, Ambrosio et al. (2025) found no significant differences in pain or disability across several osteobiologic strategies evaluated in their network meta-analysis (Ambrosio et al., 2025). Therefore, across the available evidence, radiographic fusion should not be interpreted as synonymous with superior short-term patient-reported recovery.

The available trials did not demonstrate a clear excess of major adverse events associated with bone marrow–based augmentation, but the evidence is insufficient to establish a definitive safety advantage. García de Frutos et al. (2020) reported treatment-emergent adverse events in both groups, with no events considered related to the investigational MSC product, while Chotivichit et al. (2016) reported one case of local inflammation at the marrow aspiration site (Chotivichit et al., 2016). In Pantarote and Prasopchoke (2008), donor-site pain and iliac crest fractures occurred in the ICBG group, whereas no such donor-site events were reported in the BMA group, although the difference was not statistically significant (Nipon Pantarote & Narat Prasopchoke, 2009). These findings should be interpreted cautiously because adverse-event definitions and reporting were not uniform across the included studies. Similar concerns have been identified in the broader cellular bone matrix literature, where Lambrechts et al. (2024) noted substantial variability in complication reporting and concluded that the overall evidence remained limited (Lambrechts et al., 2024). In the present review, these concerns are compounded by methodological limitations: four trials were judged to have some concerns, and three were judged to have high overall risk of bias, particularly reflecting limitations in allocation procedures or outcome assessment. Accordingly, the apparent safety of bone marrow–based augmentation should be considered reassuring but not definitively established.

Taken together, the findings indicate that bone marrow–based augmentation may be useful in selected graft constructs, but current randomized evidence does not establish it as a universally superior strategy for lumbar spinal fusion. This interpretation is broadly compatible with Gordon et al. (2024), who considered cell-based grafts a potential alternative to ICBG but highlighted heterogeneity in cell preparation and grafting, and with Ambrosio et al. (2025), whose broader network meta-analysis found that evidence for several osteobiologics remained low or very low despite favorable findings for some specific products (Ambrosio et al., 2025; Gordon et al., 2024). The main strength of the present review is its restriction to randomized trials, but important limitations include small sample sizes in several studies, heterogeneity in marrow processing and graft composition, differences in surgical techniques and fusion definitions, and variable follow-up. Future trials should therefore use adequately powered multicenter designs with standardized marrow processing and cellular characterization, clearly defined graft/scaffold protocols, standardized radiographic fusion criteria, and blinded outcome assessment. Importantly, future studies should evaluate radiographic fusion together with patient-reported outcomes and safety rather than treating fusion alone as the measure of treatment success. Overall, bone marrow–based augmentation appears promising but construct-dependent, and its superiority over conventional grafting remains unproven.

CONCLUSION

Bone marrow–based augmentation may improve radiographic fusion in selected graft constructs, but current randomized evidence does not show a consistent advantage over conventional grafting. Clinical and perioperative outcomes were generally comparable, while safety findings were limited by inconsistent reporting. Heterogeneity in marrow processing, graft composition, surgical techniques, and fusion assessment limits definitive conclusions.

REFERENCES

Ambrosio, L., Schol, J., Tamagawa, S., Muthu, S., Sakai, D., Papalia, R., Vadalà, G., & Denaro, V. (2025). Efficacy and Safety of Osteobiologics for Lumbar Spinal Fusion. Journal of Bone and Joint Surgery, 107(18), 2110–2121. https://doi.org/10.2106/JBJS.24.01205

Areesak Chotivichit, Monchai Ruangchainikom, Trongtum Tongdee, Adisak Wongkajornsilp, Parichart Permpikul, & Ekkapoj Korwutthikulrangsri. (2016). A Prospective Randomized Controlled Trial Comparing Posterolateral Lumbar Fusion With and Without Bone Marrow Concentrate Augmentation in Single-Level Lumbar Spondylolisthesis. Article in Journal of the Medical Association of Thailand, 99(10), 1073–1079. https://www.researchgate.net/publication/326089006

Chotivichit, A., Ruangchainikom, M., Tongdee, T., Wongkajornsilp, A., Permpikul, P., & Korwutthikulrangsri, E. (2016). A Prospective Randomized Controlled Trial Comparing Posterolateral Lumbar Fusion With and Without Bone Marrow Concentrate Augmentation in Single-Level Lumbar Spondylolisthesis. Journal of the Medical Association of Thailand, 99(10), 1073. http://www.jmatonline.com

García de Frutos, A., González-Tartière, P., Coll Bonet, R., Ubierna Garcés, M. T., del Arco Churruca, A., Rivas García, A., Matamalas Adrover, A., Saló Bru, G., Velazquez, J. J., Vila-Canet, G., García-Lopez, J., Vives, J., Codinach, M., Rodriguez, L., Bagó Granell, J., & Càceres Palou, E. (2020). Randomized clinical trial: expanded autologous bone marrow mesenchymal cells combined with allogeneic bone tissue, compared with autologous iliac crest graft in lumbar fusion surgery. Spine Journal, 20(12), 1899–1910. https://doi.org/10.1016/j.spinee.2020.07.014

Gordon, A., Newsome, F., Ahern, D. P., McDonnell, J. M., Cunniffe, G., & Butler, J. S. (2024). Iliac crest bone graft versus cell-based grafts to augment spinal fusion: a systematic review and meta-analysis. In European Spine Journal (Vol. 33, Number 1, pp. 253–263). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s00586-023-07941-9

Hart, R., Komzák, M., Okál, F., Náhlík, D., Jajtner, P., & Puskeiler, M. (2014). Allograft alone versus allograft with bone marrow concentrate for the healing of the instrumented posterolateral lumbar fusion. Spine Journal, 14(7), 1318–1324. https://doi.org/10.1016/j.spinee.2013.12.014

Johnson, R. G. (2014). Bone marrow concentrate with allograft equivalent to autograft in lumbar fusions. Spine, 39(9), 695–700. https://doi.org/10.1097/BRS.0000000000000254

Lambrechts, M. J., Issa, T. Z., Mazmudar, A., Lee, Y., Toci, G. R., D'Antonio, N. D., Schilken, M., Lingenfelter, K., Kepler, C. K., Schroeder, G. D., & Vaccaro, A. R. (2024). Cellular Bone Matrix in Spine Surgery – Are They Worth the Risk: A Systematic Review. In Global Spine Journal (Vol. 14, Number 3, pp. 1070–1081). SAGE Publications Ltd. https://doi.org/10.1177/21925682231205099

Nipon Pantarote, & Narat Prasopchoke. (2009). A randomized prospective study of fusion rate between bone marrow aspirate mixed local bone vs. autologous iliac crest bone graft for instrumented posterolateral fusion of lumbar spine. Lampang Medical Journal, 29(1), 12–19.

Niu, C.-C., Tsai, T.-T., Fu, T.-S., Lai, P.-L., Chen, L.-H., & Chen, W.-J. (2009). A Comparison of Posterolateral Lumbar Fusion Comparing Autograft, Autogenous Laminectomy Bone With Bone Marrow Aspirate, and Calcium Sulphate With Bone Marrow Aspirate A Prospective Randomized Study. Spine, 34(25), 2715–2719.

Park, D. K., Wind, J. J., Lansford, T., Nunley, P., Peppers, T. A., Russo, A., Hassanzadeh, H., Sembrano, J., Yoo, J., & Sales, J. (2023). Twenty-four-month interim results from a prospective, single-arm clinical trial evaluating the performance and safety of cellular bone allograft in patients undergoing lumbar spinal fusion. BMC Musculoskeletal Disorders, 24(1). https://doi.org/10.1186/s12891-023-06996-5

Pokharel, R. K., Paudel, S., & Lakhey, R. B. (2022). Iliac Crest Bone Graft Harvesting: Modified Technique for Reduction of Complications. Journal of the Nepal Medical Association, 60(247), 325–328. https://doi.org/10.31729/jnma.7086

Salamanna, F., Contartese, D., Tedesco, G., Ruffilli, A., Manzetti, M., Viroli, G., Traversari, M., Faldini, C., & Giavaresi, G. (2024). Efficacy of using autologous cells with graft substitutes for spinal fusion surgery: A systematic review and meta-analysis of clinical outcomes and imaging features. In JOR Spine (Vol. 7, Number 3). John Wiley and Sons Inc. https://doi.org/10.1002/jsp2.1347

Sayun Sumethvanich, & Anuwat Paiboon. (2023). Effectiveness of Utilizing Local Bone Graft plus Hydroxyapatite with or without Bone Marrow Aspiration from Vertebral Body in Posterolateral Lumbar Fusion: a Randomized-Controlled Trial. Lampang Medical Journal, 44(2), 48–57.