Allogenic Stem Cells in Spinal Fusion: A Systematic Review

Patrick C. Hsieh, *University of Southern California*
Zorica Buser, *University of Southern California*
Andrea C. Skelly, *Aggregate Analytics, Inc*
Erica D. Brodt, *Aggregate Analytics, Inc*
Darrel Brodke, *University of Utah*
Hans-Joerg Meisel, *Bergmannstrost Hospital*
Jong-Boem Park, *Emory University*
Sangwook Yoon, *Emory University*
Jeffrey C. Wang, *University of Southern California*

Final published version: http://dx.doi.org/10.1177/2192568219833336

Accessed November 28, 2021 3:03 AM EST

knowledge FORUM
AOSPINE

Allogenic Stem Cells in Spinal Fusion: A Systematic Review

Patrick C. Hsieh, MD[1], Zorica Buser, PhD[1], Andrea C. Skelly, PhD[2], Erika D. Brodt[2], Darrel Brodke, MD[3], Hans-Joerg Meisel, MD, PhD[4], Jong-Beom Park, MD[6], S. Tim Yoon, MD, PhD[6], and Jeffrey C. Wang, MD[1]

Abstract

Study Design: Systematic review.

Objectives: To review, critically appraise, and synthesize evidence on the use of allogenic stem cell products for spine fusion compared with other bone graft materials.

Methods: Systematic searches of PubMed/MEDLINE, through October 31, 2018 and of EMBASE and ClinicalTrials.gov through April 13, 2018 were conducted for literature comparing allogenic stem cell sources for fusion in the lumbar or cervical spine with other fusion methods. In the absence of comparative studies, case series of ≥10 patients were considered.

Results: From 382 potentially relevant citations identified, 6 publications on lumbar fusion and 5 on cervical fusion met the inclusion criteria. For lumbar arthrodesis, mean Oswestry Disability Index (ODI), visual analogue scale (VAS) pain score, and fusion rates were similar for anterior lumbar interbody fusion (ALIF) using allogenic multipotent adult progenitor cells (Map3) versus recombinant human bone morphogenetic protein–2 (rhBMP-2) in the one comparative lumbar study (90% vs 92%). Across case series of allogenic stem cell products, function and pain were improved relative to baseline and fusion occurred in ≥90% of patients at ≥12 months. For cervical arthrodesis across case series, stem cell products improved function and pain compared with baseline at various time frames. In a retrospective cohort study fusion rates were not statistically different for Osteocel compared with Vertigraft allograft (88% vs 95%). Fusion rates varied across time frames and intervention products in case series.

Conclusions: The overall quality (strength) of evidence of effectiveness and safety of allogenic stem cells products for lumbar and cervical arthrodesis was very low, meaning that we have very little confidence that the effects seen are reflective of the true effects.

Keywords

fusion, allogenic cells, stem cells, systematic review

Introduction

Spine fusions is an accepted and commonly performed procedure for treatment of various spinal conditions, including degenerative disc disease, degenerative spinal deformity, traumatic spine fractures, and spinal tumors. Achieving solid bony arthrodesis is one of the primary goals of spine fusion surgery. Without it, patients can have poor postsurgical outcomes, develop painful pseudoarthrosis, and may require revision surgeries.[1-4] Nevertheless, achieving solid arthrodesis in complex spine surgeries can be a daunting task as pseudoarthrosis rate for adult spinal deformity surgery can range from 0% to 25%[2,3,5-8] and pseudoarthrosis for spine tumors can range from 0% to 43%.[9-12]

[1] University of Southern California, Los Angeles, CA, USA
[2] Aggregate Analytics, Inc, Fircrest, WA, USA
[3] University of Utah School of Medicine, Salt Lake City, UT, USA
[4] Bergmannstrost Hospital, Halle, Germany
[5] Uijongbu St Mary's Hospital, The Catholic University of Korea School of Medicine, Uijongbu, Korea
[6] Emory University, Atlanta, GA, USA

Corresponding Author:
Patrick C. Hsieh, Department of Neurological Surgery, University of Southern California Keck School of Medicine, 1200 North State Street, Suite 3300, Los Angeles, CA 90033, USA.
Email: phsieh@usc.edu

The use of autologous iliac crest bone graft (ICBG) has been considered the "gold standard" for spine fusion.[13-15] ICBG is believed to possess the 3 essential properties of osteogenesis, osteoinductivity, and osteoconductivity to optimize spine arthrodesis. However, ICBG may have limited quantity and quality dependent on the age and biology of the patients.[14,16-18] Moreover, harvest of ICBG is also associated with increased rate of morbidities such as chronic pain at the harvest site, infection, and hematoma.[14,15,19-21] As a result, surgeons and researchers are constantly seeking for the best substitute for ICBG that can help achieve bony arthrodesis in complex spine fusions.

Over the past couple of decades, commonly used alternative bone graft substitutes includes allograft, demineralized bone matrix, ceramics, synthetic bone substitutes, and bone morphogenetic proteins. In recent years, there has been a rising interest in the use of stem cells products as bone graft substitute for spine fusions and other orthopedic procedures.[22-28] Stem cells or progenitor cells are undifferentiated cells that are uncommitted in their cellular differentiation that also possess the property of self-renewal, and they can be driven toward specific tissue or cellular differentiation with growth factors and mitogens.[16,17,29-32] For spine procedures, mesenchymal stem cells and osteogenic progenitor cells may be differentiated into osteoblasts that can help achieve fusion by providing cells that participate in bone formation and may also produce osteoinductive molecules.[16,17,29-32] They can be isolated from autologous and allogeneic sources then expanded to provide a constant and viable source of bone graft substitute alternative to ICBG. Unlike other bone graft substitutes available in the market today, mesenchymal stem cells and progenitor cells are believed to possess the osteogenic and osteoinductive properties of ICBG.[17,32] In addition, when mix with other bone graft extenders as carriers or fillers, they can add osteoconductivity to the final stem cell products.[17,32] While the potential of mesenchymal stem cells and osteogenic progenitor cells to provide the clinical results in spine fusion as ICBG is quite appealing, we need supportive data to confirm that assertion. The aim of this study was to systematically review, critically appraise, and synthesize evidence on use of stem cells from allogenic bone marrow aspirate, adipose, or any other allogenic sources for spine fusion compared with other graft materials.

Key Questions

With regard to use of allogenic stem cells for arthrodesis (fusion) in adult patients with degenerative conditions of the lumbar or cervical spine:

Key Question 1: Is use of allogenic stem cells for fusion as effective as fusion with standard autograft or other graft materials in the thoracolumbar or cervical spine?

Key Question 2: What complications are associated with allogenic stem cell use in fusion? Is use of stem cells safer than fusion with standard allograft or autograft in the thoracolumbar or cervical spine?

Key Question 3: Is there evidence that patient factors (eg, age, smoking, comorbid conditions, revision status, presence of deformity), number of levels treated, cell type, or preparation modify the association between allogenic stem cell–based fusion and the primary outcomes?

Key Question 4: Is allogenic stem cell use for fusion cost-effective compared with other graft materials?

Figure 1 provides an overview of the patients, interventions and outcomes considered for these questions.

Materials and Methods

The methods for this systematic review followed accepted standards for systematic review/comparative effectiveness reviews for rigor, quality, and transparency, including those described by the Agency for Healthcare Research and Quality (AHRQ), Institute of Medicine (IOM) Standards for Systematic Reviews, and the Patient-Centered Outcomes Research Institute (PCORI) Methods Guide.[33-35]

Electronic Literature Search

Systematic searches of PubMed/MEDLINE, through October 31, 2018 and EMBASE and ClinicalTrials.gov through April 13, 2018 were conducted. Only studies with abstracts in humans, written in English were considered for inclusion, with no other limits were placed on the search. A priori inclusion and exclusion criteria are detailed in Appendix A (available in the online supplement of the journal). Briefly we sought to identify comparative studies of allogenic stem cells use versus more commonly used methods of fusion (eg, autograft) in persons with degenerative disease of the cervical or thoracolumbar spine. In the absence of comparative studies, case series were considered. The search strategy included use of controlled vocabulary (MeSH terms) as well as keywords. (Appendix A). Bibliographies of included studies and relevant systematic reviews were reviewed to identify pertinent studies. Citations were dual reviewed for inclusion at both title/abstract and full-text stages. ClincalTrials.gov was searched to identify studies which may have new publications. (Appendix G).

Data Extraction

In addition to results for primary and secondary outcomes, data abstraction included patient characteristics, demographics, lifestyle factors (eg, smoking), comorbidities (eg, obesity), cointerventions (eg, pharmaceutical, physical therapy), intervention, and comparator details (eg, spinal levels treated, cell preparation and concentration, delivery) if reported.

Study Quality

Each included study was independently assessed for risk of bias and methodological quality by 2 reviewers (ACS, AF) using pre-set criteria based on methods delineated in the *Cochrane*
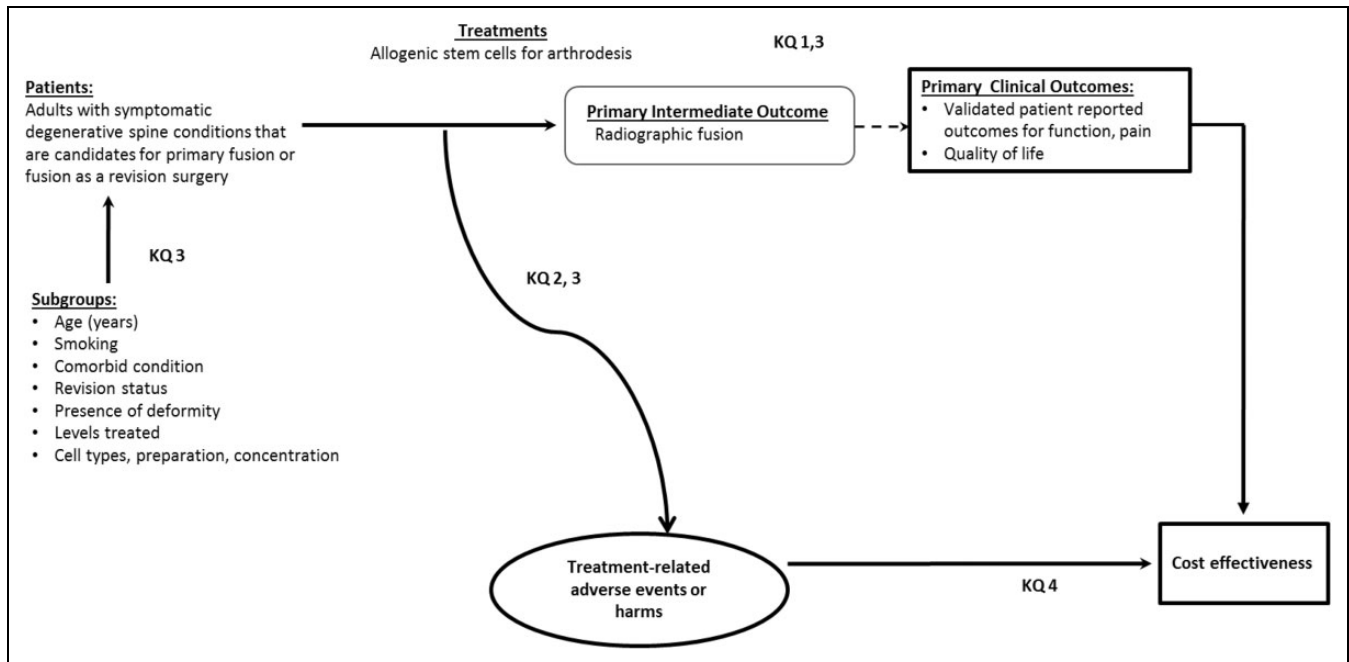
Figure 1. Analytic framework.

Handbook for Systematic Reviews of Interventions, *The Journal of Bone and Joint Surgery,* and the Agency for Healthcare Research and Quality with adaptations focusing on criteria associated with methodological quality (Appendix C).[33,36,37] Economic studies were evaluated according to The Quality of Health Economic Studies (QHES) instrument developed by Ofman et al.[38] Where feasible, the focus was on comparative studies with the least potential for bias and the fewest limitations. Risks of bias assessments are detailed in Appendix C.

Data Analysis

Statistical testing was generally not performed for observational studies except for complications. Risk ratios were calculated for dichotomous outcomes for complications from comparative studies if differences between groups were or approached statistical significance using the Rothman EpiSheet; for continuous outcomes, mean differences were calculated and 2-sample t tests if appropriate using GraphPad.[39,40] Study design, heterogeneity across studies and variation in reporting precluded the pooling of data.

Overall Strength of Body of Evidence

For function, pain, fusion and adverse events, the overall strength of evidence across included studies was assessed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group tenets and recommendations made by the AHRQ and is further described in Appendix D.[33,41-43] The overall quality of evidence was based on studies at least risk for bias. In determining the quality

(strength) of a body of evidence regarding a given outcome, the overall quality may be downgraded 1 or 2 levels based on the following domains: (1) risk of bias due to study limitations, (2) consistency (heterogeneity) of results, (3) directness of evidence (eg, hard clinical outcomes), (4) precision of effect size estimates (eg, width of confidence intervals), and (5) publication or reporting bias. Publication and reporting bias are difficult to assess, particularly with fewer than 10 randomized controlled trials.[33] Publication bias was unknown in all studies and thus this domain was eliminated from the strength of evidence tables. The initial quality of the overall body of evidence begins as *high* for randomized controlled trials and *low* for observational studies. The body of evidence for methodologically strong observational studies may be upgraded 1 or 2 levels if there are no downgrades in the primary domains above *and* one or more of the following are met: (1) large magnitude of effect, (2) dose-response gradient, and (3) all plausible biases would decrease the magnitude of an apparent effect. The final overall quality (strength) of the body of literature expresses the confidence in the estimate of effect and the impact that further research may have on the results as follows:

- *High:* High confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect.
- *Moderate:* Moderate confidence that the evidence reflects the true effect. Further research may change our confidence in the estimate of effect and may change the estimate.
- *Low:* Low confidence that the evidence reflects the true effect. Further research is likely to change the confidence in the estimate of effect and likely to change the estimate.
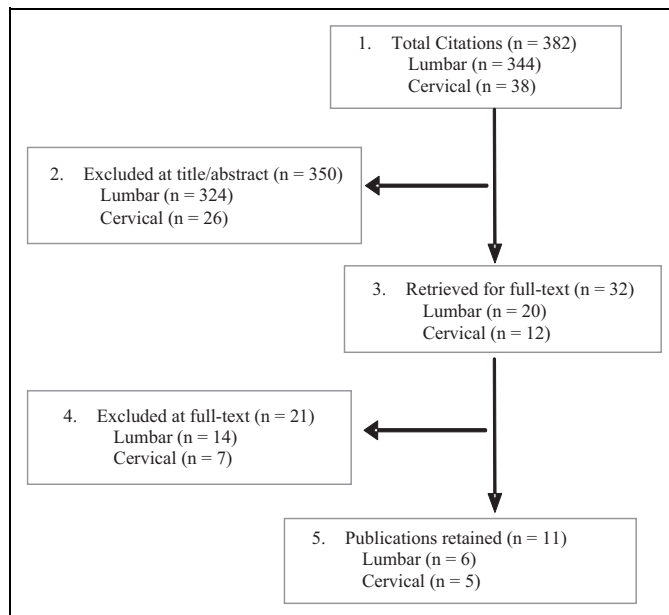
1. Total Citations (n = 382)
   Lumbar (n = 344)
   Cervical (n = 38)

2. Excluded at title/abstract (n = 350)
   Lumbar (n = 324)
   Cervical (n = 26)

3. Retrieved for full-text (n = 32)
   Lumbar (n = 20)
   Cervical (n = 12)

4. Excluded at full-text (n = 21)
   Lumbar (n = 14)
   Cervical (n = 7)

5. Publications retained (n = 11)
   Lumbar (n = 6)
   Cervical (n = 5)

Figure 2. Flowchart showing results of literature search.

- *Very low:* Very little confidence that the evidence reflects the true effect; the true effect is likely to be substantially different than the estimated effect. In addition, this rating may be used if there is no evidence or it is not possible to estimate an effect.

Results

Study Selection and Risk of Bias

From 382 potentially relevant citations identified, eleven studies (from all sources) met the inclusion criteria (Figure 2). Studies excluded at full-text review are listed in Appendix B. Nine ongoing clinical trials were identified (Appendix G). The 2 retrospective cohort studies, one on lumbar fusion[44] and one on cervical fusion[45] were identified and were considered at moderately high risk of bias. Methodological concerns included selection bias, lack of blinded outcomes assessment and failure to control for confounding. The 9 case series were considered at high risk of bias (Appendix C). Across studies, there was substantial heterogeneity with regard to stem cell products and preparations used. This, combined with poor study quality, precluded pooled analysis. No studies formally evaluated differential efficacy or safety. No full economic studies were identified.

Key Question 1: Effectiveness of Allogenic Cells for Arthrodesis

Lumbar Arthrodesis. Six studies of lumbar fusion met the inclusion criteria including one retrospective cohort comparing allogenic multipotent adult progenitor cells (Map3) versus recombinant human bone morphogenetic protein–2 (rhBMP-2),[44] 3 case series

using Osteocel Plus (mesenchymal stem cells with demineralized bone matrix) alone[46,47] or in combination with local autograft,[48] one case series of using Trinity Evolution cellular bone allograft combined with local autograft,[49] and one case series of VIA Graft cellular bone matrix.[50] Across studies, mean ages were 50 to 64 years and 37% to 53% of participants were male. Diabetes was a common comorbidity ($\sim$20% of participants). Smoking frequency ranged from 9% to 42%. Surgical approaches varied across studies. Patient and intervention characteristics are summarized in Table 1. Details of patient populations and procedures are provided in Appendix F.

The small retrospective cohort (N = 41) comparing anterior lumbar interbody fusion (ALIF) using Map3 with ALIF using rhBMP-2[44] found no difference between them for function based on Oswestry Disability Index (ODI) or in visual analogue scale (VAS) pain at 12 months. (Table 2) Across levels, both treatments yielded similar fusion rates (90% vs 92%). While some differences in fusion rate were noted based on the number of levels fused, there was no distinct pattern favoring one treatment over the other; sample size was however small.

Only one small series of XLIF using Osteocel Plus (N = 35) provided data on both function and pain, reporting improvement in both at 12 months relative to baseline.[47] Similarly, the case series of posterolateral fusion (PLF) using Trinity Evolution with local autograft (N = 43) reported significant improvement in both VAS back pain and leg pain at 12 months compared with baseline[49] (Table 3).

One small series of Osteoplus[47] (N = 35) found 56% improvement from baseline in total Short Form–36 (SF-36) scores at 12 months and reported that the majority of patients were at least somewhat satisfied with their outcomes (Appendix E).

Across 5 case series of various allogenic cell products and fusion approaches (N = 158), fusion occurred in $\geq$90% of patients and levels by 12 to 14 months[46-50] (Table 3).

Cervical Arthrodesis. Five studies of cervical fusion in patients with degenerative disease met the inclusion criteria, including 1 retrospective cohort comparing Osteocel with allograft and 3 case series. The case series include 1 case series using Osteocel Plus,[51] 2 case series using Trinity Evolution with local autograft,[52,53] and 1 using Vivigen allograft.[54] Anterior cervical discectomy and fusion (ACDF) was performed in 3 of the 4 series; a variety of procedures was used in the fourth case series. Across studies, ages ranged from 47 to 64 years and proportion of males ranged from 28% to 56%. Reported tobacco use ranged from 16% to 37% with 1 series reporting 45% of patients were current or former smokers (Table 4).

Across the case series, functional improvement based on neck disability index (NDI) and improvement in VAS pain relative to baseline were reported at multiple time frames. Two series (N = 69) using Trinity Evolution reported significant improvement in NDI and VAS neck pain at 6 and 12 months.[52,53]

A small series (N = 21) using Vivigen reported an 85.9% decreased in NDI from baseline and 81.4% decrease in VAS

Table 1. Summary of Patient Demographics and Intervention Characteristics for Studies Evaluating the Use of Allogenic Stem Cells for Lumbar Fusion

Cohort Study, RoB	Demographics (Mean or %)	Diagnosis	Intervention Characteristics, Previous Treatment	Treatment Groups	Levels, %
Lee 2017 *Moderately high RoB*	N = 41, 66 segments Age: 53 years Male: 39% Smokers: 20% Symptom duration: NR	Lumbar DDD	ALIF Failed conservative treatment: NR Prior surgery: 0% at operative level (study excluded those with prior failed fusion at same operative level)	Anterior standalone interbody device packed with: • <u>Group 1:</u> Map3 allograft • <u>Group 2:</u> rhBMP-2 allograft	1, 2, or 3 levels Levels treated (%): NR

Case Series, RoB	Demographics (Mean or %)	Diagnosis	Intervention Characteristics, Previous Treatment	Graft Materials	Levels, %
Posterior approach					
Musante 2016 *High RoB*	N = 43 Age: 65 years Male: 37% Smokers: 9% Symptom duration: NR	Lumbar DDD (with any of the following: radiculopathy, stenosis, spondylolisthesis, HNP, facet joint degeneration)	PLF with decompressive laminectomy Failed conservative treatment: NR Prior surgery: NR	Trinity Evolution allograft + local autograft bone	1 or 2 level (L1 to S1) • 1-level: NR • 2-level: NR • Levels treated (%): NR
Multiple approaches					
Kerr 2011 *High RoB*	N = 52 Age: 50 years Male: 40.4% Smokers: 42% Symptom duration: NR	DDD, HNP, or spondylolisthesis	Circumferential fusion (67%), TLIF (16%), or ALIF (17%) Failed conservative treatment: NR Prior surgery: 21%	Osteocel allograft	1 or 2 Level • 1-level: 69% • 2-level: 31%
Tohmeh 2012 *High RoB*	N = 78 Age: 52 years Male: 58% Smokers: 13% Symptom duration: NR	Lumbar DDD, spondylolisthesis, postlaminectomy syndrome, Adjacent segment disease, scoliosis, retrolisthesis, facet disease, stenosis, or fracture	XLIF or TLIF Failed conservative treatment: 100% Prior surgery: 65%	Osteocel allograft	Levels treated (%): NR • 1, 2, or 3 levels: NR • Levels treated (%): • L1/2: 4% • L2/3: 16% • L3/4: 26% • L4/5: 43% • L5/S1: 10%
Minimally invasive transforaminal approach					
Ammerman 2013 *High RoB*	N = 23, 26 levels Age: 57 years Male: 52% Smokers: 4% Symptom duration: NR	Degenerative lumbar disease with or without radiculopathy	MITLIF Failed conservative treatment: 100% (minimum of 3 months) Prior surgery: 0%	Osteocel Plus allograft + PEEK interbody spacer + local autograft	1 or 2-level • 1-level: 87% • 2-level: 13% • Levels treated (%): NR
Tally 2018 *High RoB*	N = 75, 85 levels Age: 58 years Male: 53% Smokers: 21% Symptom duration: NR	Degenerative conditions (radiculopathy and any of the following: retrolisthesis, facet cyst, adjacent segment, stenosis, spondylolysis, spondylolisthesis, disk herniation, scoliosis)	MITLIF Failed conservative treatment: NR Prior surgery: NR	VIA Allograft + β-TCP + concentrated BMA	1 or 2-Level • 1-level: 87% • 2-level: 13% • Levels treated (%): NR

Abbreviations: β-TCP, β-tricalcium phosphate; ALIF, anterior lumbar interbody fusion; BMA, bone marrow aspirate; DDD, degenerative disc disease; HNP, herniated nucleus pulposus; Map3, cellular allogenic bone graft, which includes cortical cancellous bone chips, demineralized bone matrix, and cryogenically preserved, viable multipotent adult progenitor cells; MITLIF, minimally invasive transforaminal lumbar interbody fusion; NR, not reported; rhBMP-2, recombinant human bone morphogenic protein–2; RoB, risk of bias; PEEK, polyetheretherketone; PLF, posterolateral fusion; TLIF, transforaminal lumbar interbody fusion; XLIF, extreme lateral interbody fusion.

Table 2. Summary of Nonrandomized, Retrospective Comparative Data (Lee, 2017): Allogenic Multipotent Adult Progenitor Cells (Map3) for Lumbar Fusion

Lee 2017 Retrospective Cohort, Moderately High RoB, Outcome Measure	ALIF With Map3 Mean ± SD (N) or % (n/N)	ALIF With rhBMP-2 Mean ± SD (N) or % (n/N)	P[a]
Effectiveness data			
ODI (0-100, higher score = worse disability)			
Baseline	57.8 ± 6.1 (n = 20)	57.0 ± 6.4 (n = 21)	N/A
12 months	24.8 ± 7.2 (n = 20)	25.7 ± 9.6 (n = 21)	.966
VAS (0-10, higher score = worse pain)			
Baseline	7.6 ± 1.1 (n = 20)	7.7 ± 0.8 (n = 21)	N/A
12 months	2.2 ± 1.1 (n = 20)	2.6 ± 1.1 (n = 21)	.251
Fusion at 12 months (from radiograph)[b]			
Overall	90% (27/30 levels)	92% (33/36 levels)	.890
1 level	100% (11/11 levels)	78% (7/9 levels)	NR
2 levels	81% (13/16 levels)	100% (18/18 levels)	NR
3 levels	100% (3/3 levels)	89% (8/9 levels)	NR
Safety Data			

Complications			Effect Size[c] P
Postoperative complications, any[d]	30% (6/20) (6 events)	57% (12/21) (16 events)	RR 0.53 (95% CI 0.24, 1.13); P = .060
Posterior infection, drainage	15% (3/20)	19% (4/21)	NR
Postoperative radiculitis	10% (2/20)	38% (8/21)	NR
Epidural hematoma	5% (1/20)	5% (1/21)	NR
Retroperitoneal hematoma	0% (0/20)	10% (2/21)	NR
Partial right foot drop	0% (0/20)	5% (1/21)	NR
Nonunion	10% (3/30 levels)	8% (3/36 levels)	NR
1 level	0% (0/11 levels)	22% (2/9 levels)	NR
2 levels	19% (3/16 levels)	0% (0/18 levels)	NR
3 levels	0% (0/3 levels)	11% (1/9 levels)	NR
Death	0% (0/20)	0% (0/21)	N/A

Abbreviations: ALIF, anterior lumbar interbody fusion; CI, confidence interval; Map3 = cellular allogenic bone graft which includes cortical cancellous bone chips, demineralized bone matrix, and cryogenically preserved, viable multipotent adult progenitor cells; N/A, not applicable; NR, not reported; ODI, Oswestry Disability Index; rhBMP-2, recombinant human bone morphogenetic protein–2; RR, risk ratio; SD, standard deviation; VAS, visual analogue scale.
[a]As reported by authors for comparison between groups.
[b]Fusion was defined as evidence of bridging across endplates or bridging from endplates to interspace disc plugs.
[c]RR calculated by Aggregate Analytics, Inc. (AAI) using the Rothman Episheet.
[d]Patients could have more than 1 event.

pain at 12 months.[54] The largest series, using Osteocel (N = 135 patients with 24-month follow-up) reported significant improvement at 24 months[51] for both NDI and VAS neck pain and arm pain. However, the authors did not provide data for verification (Table 5). In addition, SF-12/36 physical component summary (PCS) scores improved relative to baseline at 6 and 12 months in one series using Trinity Evolution[52] and at 24 months in one series of Osteocel Plus.[47] (Appendix E).

In the retrospective cohort study (N = 114), overall fusion rates were somewhat lower in the Osteocel group compared with the Verigraft allograft controls (88% vs 95%), but statistical significance was not achieved; a similar pattern was seen for those receiving intervention at one level (86.2% vs 96.6%) only. Fusion between the treatment groups was similar for 2-level procedures (Table 6). Authors do not report measures of function or pain.

Across case series, criteria and definitions of fusion varied and fusion frequency varied across time frames and intervention products. At 6 months, they varied with rates in 2 series of Trinity Evolution[52,53] of 66% and 79%, respectively, and of 100% in the Vivigen series[54]; by 12 months, fusion was seen in 89% and 94% of Trinity Evolution recipients. At 24 months, the series of Osteocel reported fusion frequency of 87%[51] (Table 5).

Key Question 2: Safety, Adverse Events, Harms

Adverse events and harms were poorly reported across studies of both lumbar and cervical arthrodesis and sample sizes were likely inadequate to detect rare events; details are provided in Appendices E and F.

Lumbar Arthrodesis. The small retrospective cohort (N = 41) reported that complications overall were less common

Table 3. Summary of Case Series Outcomes: Allogenic Cells for Lumbar Fusion.

Outcome Measure F/U (Months)	Allograft Cell/Intervention	Author, Year	Mean ± SD (N), or % (n/N)	P Value Based on Δ vs Baseline[a]
ODI[b]				
Baseline	XLIF, Osteocel Plus	Tohmeh, 2012	45.7 ± NR (n = 35)	N/A
12 months			27.1 ± NR (n = 35)	<.05
			41% improvement	
VAS back pain[b]				
Baseline	PLF, Trinity Evolution + local autograft	Musante, 2016	6.3 ± NR (n = 43)	N/A
	XLIF, Osteocel Plus	Tohmeh, 2012	7.4 ± NR (n = 35)	N/A
3 months	PLF, Trinity Evolution + local autograft	Musante, 2016	2.0 ± NR (n = 43)	<.0001
12 months	PLF, Trinity Evolution + local autograft	Musante, 2016	2.1 ± NR (n = 43)	<.0001
	XLIF, Osteocel Plus	Tohmeh, 2012	3.4 ± NR (n = 35)	<.05
			55% improvement	
VAS leg pain[b]				
Baseline	PLF with Trinity Evolution	Musante, 2016	7.1 ± NR (n = 43)	N/A
	XLIF with Osteocel Plus	Tohmeh, 2012	6.8 ± NR (n = 35)	N/A
3 months	PLF, Trinity Evolution + local autograft	Musante, 2016	1.5 ± NR (n = 43)	<.0001
12 months	PLF, Trinity Evolution + local autograft	Musante, 2016	1.8 ± NR (n = 43)	<0.0001
	XLIF, Osteocel Plus	Tohmeh, 2012	3.8 ± NR (n = 35)	<.05
			43% improvement	
Fusion				
12 months	MITLIF, Osteocel Plus + autograft (facet)	Ammerman, 2013	Overall: 91% (21/23)	N/A
	XLIF, Osteocel Plus	Tohmeh, 2012	90% of levels (55/61 levels)	N/A
	PLF, Trinity Evolution cellular bone allograft + local autograft	Musante, 2016	91% (39/43)	N/A
	MITLIF, VIA Graft Cancellous Sponge soaked in concentrated BMA	Tally, 2018	96% (72/75)	N/A
Median 14 months (range 8-27 months)	Circumferential (67%), ALIF (17%) or TLIF (16%) with Osetocel Plus	Kerr, 2011	92% (48/52)	N/A

Abbreviations: ALIF, anterior lumbar interbody fusion; F/U, follow-up; ICBG, iliac crest bone graft; MITLIF, minimally invasive transforaminal lumbar interbody fusion; N/A, not applicable; NR, not reported; ODI, Oswestry Disability Index; PLF, posterolateral fusion; SD, standard deviation; SF-36, Short Form–36 questionnaire; VAS, visual analogue scale; XLIF, extreme lateral interbody fusion.
[a]As reported by authors.
[b]ODI is rated on a scale of 0 to 100 with higher scores indicating worse disability. VAS was rated on a scale of 0 to 10 with higher scores indicated worse pain.

following ALIF using Map3 (30%) compared with rhBMP-2[44] (57%), but statistical significance was not reached due in part to the small number of patients (Table 2). Postoperative radiculitis in particular was less common in Map3 versus rhBMP-2 recipients (10% vs 38%). No differences in nonunion, posterior infection or epidural hematoma were seen. Across the 4 case series (Appendices E and F), nonunion was reported in 8% to 9% of patients.[46-49] Reporting of other complications was sparse. No intraoperative complications were seen in three studies[47,49,50] and none related to stem cell products were seen in 2 series[48,49]; however, small samples sizes limit evaluation of complications particularly for rare events. Additional surgical procedures were done in 7% of patients in one series[49] and may not be related to stem cell product use. Pseudarthrosis leading to revision occurred in 4% (3/75) in one series.[50] Surgical intervention at an adjacent level (1 patient) was reported in another small series (N = 40).[47]

Cervical Arthrodesis. The comparative retrospective cohort study reported more nonunion for Map3 recipients compared with allograft recipients (12.3% vs 5.3%) at 12 months, but results did not reach statistical significance.[45] Among patients with failed fusion 4/7 in the Osteocel group versus 3/3 in the allograft group required revision. No treatment-related adverse events were reported, but small sample size may preclude identification of rare events (Table 6). Across case series, rates of nonunion varied substantially (Appendix E). At 6 months, the 2 series of Trinity Evolution reported nonunion rates of 34.4%[52] and 21.4%,[53] with lower rates by 12 months (10.6% and 6.5%). While the small series of Vivigen reported no nonunions,[54] the series of Osteocel reported 18% at 24 months with 13% at >24 months.[51] No revision surgeries were reported across 3 series.[51-53]

Key Questions 3 (Modification of Treatment Effect) and 4 (Economic Studies)

Included studies were not designed to evaluation modification of treatment effect by demographic or other factors. Subgroup analysis in one small lumbar arthrodesis case series found no difference in back pain, leg pain, or fusion frequency between those <65 years old and those ≥65 years old[49]; however, firm conclusions are not possible given the potential for high risk of

Table 4. Summary of Patient Demographics and Intervention Characteristics for Comparative Studies and Case Series Evaluating the Use of Allogenic Stem Cells for Cervical Fusion.

Cohort Studies, RoB	Demographics (Mean or %)	Diagnosis	Intervention Characteristics, Previous treatment	Treatment Groups	Levels, %
McAnany 2015 *Moderately high RoB*	N = 114 Age: 47 years Male: 56% Smokers: 37% Symptom duration: NR	Cervical DDD; radiculopathy or myelopathy	ACDF Failed conservative treatment: NR Prior surgery: NR	1 of 2 allografts: • Group 1: Osteocel • Group 2: Vertigraft	1- or 2-Level • 1-level: 50.9% • 2-level: 49.1% Levels treated (%): NR
Case Series, RoB	**Demographics (Mean or %)**	**Diagnosis**	**Intervention Characteristics, Previous Treatment**	**Graft Materials**	**Levels, %**
Anterior approach					
Eastlack 2014 *High RoB*	N = 182, 249 levels Age: 51 years Male: 51% Smokers: 25% Symptom duration: NR	Cervical DDD	ACDF Failed conservative treatment: 100% Prior surgery: 0% at operative levels (study excluded those with previous surgery at same operative levels)	Osteocel Plus allograft + PEEK interbody spacer	1- or 2-Level (C3-T1) • 1-level: 62% • 2-level: 38% Levels treated (%): • C3/4: 7% • C4/5: 16% • C5/6: 64% • C6/7: 47% • C7/T1: 2%
Peppers 2017 *High RoB*	N = 40, 80 levels Age: 49 years Male: 28% Smokers: 45% Symptom duration: NR	Cervical DDD	ACDF Failed conservative treatment: NR Prior surgery: 0%	Trinity Evolution allograft + PEEK interbody spacer	2-Level: 100% Levels treated (%): • C5/6/7: 65% • C4/5/6: 32.5% • C3/4/5: 2.5%
Vanichkachorn 2016 *High RoB*	N = 31 Age: 49 years Male: 39% Smokers: 16% Symptom duration: NR	Cervical DDD	ACDF Failed conservative treatment: NR Prior surgery: 9.7%	Trinity Evolution allograft + PEEK interbody spacer	1-Level: 100% Levels treated (%): • C3/4: 3.2% • C4/5: 12.9% • C5/6: 51.6% • C6/7: 32.3%
Multiple approaches					
Divi 2017 *High RoB*	N = 21 Age: 64 years Male: 33% Smokers: 24% Symptom duration: NR	Cervical myelopathy, CSM, HNP, OPLL, cervical radiculopathy, C1/C2 instability (with rheumatoid arthritis, unstable fractures, or prior attempted odontoid fixation with nonunion)	ACDF (38%), ACCF (14%), or PCF (48%) Failed conservative treatment: 100% Prior surgery: NR	Vivigen allograft + interbody spacer	NR

Abbreviations: ACCF, anterior cervical corpectomy and fusion; ACDF, anterior cervical discectomy and fusion; CSM, cervical spondylotic myelopathy; HNP, HNP, herniated nucleus pulposus; NR, not reported; PCF, posterior cervical fusion; PEEK, polyetheretherketone; RoB, risk of bias; OPLL, ossification of posterior longitudinal ligament.

bias in this study and lack of comparison with another treatment (Appendices E and F).

No full economic studies comparing allogenic stem cells products with other methods of fusion were identified.

Evidence Summary, Overall Quality (Strength) of Evidence

For both lumbar and cervical arthrodesis, the overall quality of evidence is very low with regard to the benefits and safety of allogenic cell sources for fusion, meaning we have very little confidence that the effects seen reflect the true effects (Tables 7 and 8). Case series comprise the bulk of the evidence. Only 2 cohort studies, one evaluating lumbar arthrodesis and the other cervical arthrodesis were identified. High risk of bias, lack of precision due to small sample sizes, and inability to evaluate consistency across studies resulted in grading the overall evidence as very low.

Table 5. Summary of Function, Pain, and Fusion Outcomes From Case Series: Allogenic Cells for Cervical Fusion.

Outcome Measure F/U (Months)	Approach/Allograft	Author, Year	Mean $\pm$ SD (n)	P Value Based on Δ vs Baseline[a]
Function[b]				
NDI[c]				
Baseline	ACDF, Trinity Evolution	Peppers, 2017	44 $\pm$ NR (n = 40)	N/A
	ACDF, Trinity Evolution	Vanichkachorn, 2016	45 $\pm$ NR (n = 31)	N/A
	ACDF, ACCF or PCF; Vivigen	Divi, 2017	40.3 $\pm$ NR (n = 21)	N/A
6 months	ACDF, Trinity Evolution	Peppers, 2017	23 $\pm$ NR (n = 35)	<.0001
	ACDF, Trinity Evolution	Vanichkachorn, 2016	21 $\pm$ NR (n = 28)	<.05
12 months	ACDF, Trinity Evolution	Peppers, 2017	19 $\pm$ NR (n=38)	<.0001
	ACDF, Trinity Evolution	Vanichkachorn, 2016	17 $\pm$ NR (n = 31)	<.05
	ACDF, ACCF, or PCF; Vivigen	Divi, 2017	6 $\pm$ NR (n = 21)	NR
Mean Δ from baseline NDI[c]				
24 months	ACDF, Osteocel	Eastlack, 2014[d]	−21.2 $\pm$ NR (n = 135)	<.001[e]
Average % Decrease in NDI[c]				
12 months	ACDF, ACCF or PCF; Vivigen	Divi, 2017	85.9% (n = 21)	NR
Pain[b]				
VAS Neck[c]				
Baseline	ACDF, Trinity Evolution	Peppers, 2017	59 $\pm$ NR (n = 40)	N/A
	ACDF, Trinity Evolution	Vanichkachorn, 2016	59 $\pm$ NR (n = 31)	N/A
	ACDF, ACCF or PCF; Vivigen	Divi, 2017	83 $\pm$ NR (n = 21)	N/A
6 months	ACDF, Trinity Evolution	Peppers, 2017	18 $\pm$ NR (n = 35)	<.0001
	ACDF, Trinity Evolution	Vanichkachorn, 2016	14 $\pm$ NR (n = 28)	<.05
12 months	ACDF, Trinity Evolution	Peppers, 2017	17 $\pm$ NR (n = 38)	<.0001
	ACDF, Trinity Evolution	Vanichkachorn, 2016	11 $\pm$ NR (n = 31)	<.05
Mean Δ from baseline VAS Neck[c]				
24 months	ACDF, Osteocel	Eastlack, 2014[d]	−34 $\pm$ NR (n = 135)	<.001
Average % Decrease in VAS[c]				
12 months	ACDF, ACCF or PCF; Vivigen	Divi, 2017	81.4% (n = 21)	NR
Mean Δ from baseline VAS arm[c]				
24 months	ACDF, Osteocel	Eastlack, 2014[d]	−35 $\pm$ NR (n = 135)	<.001
VAS Left/Right arm[c]				
Baseline	ACDF, Trinity Evolution	Vanichkachorn, 2016	Left: 44 $\pm$ NR (n = 31) Right: 36 $\pm$ NR (n = 31)	N/A
6 months			Left: 13 $\pm$ NR (n = 28) Right: 6 $\pm$ NR (n = 28)	Left: <.05 Right: <.05
12 months			Left: 12 $\pm$ NR (n = 31) Right: 4 $\pm$ NR (n = 31)	Left: <.05 Right: <.05
Fusion			% of patients (n/N)	
6 months	ACDF, Trinity Evolution	Peppers, 2017	65.7% (23/35)	N/A
	ACDF, Trinity Evolution	Vanichkachorn, 2016	78.6% (22/28)	N/A
	ACDF, ACCF or PCF; Vivigen	Divi, 2017	100% (21/21)	N/A
12 months	ACDF with Trinity Evolution	Peppers, 2017	89.4% (34/38)	N/A
	ACDF with Trinity Evolution	Vanichkachorn, 2016	93.5% (29/31)	N/A
24 months[e]	ACDF with Osteocel	Eastlack, 2014	87% (157/180)	N/A

Abbreviations: ACDF, anterior cervical discectomy and fusion; F/U, follow-up; MSCs, mesenchymal stem cells; N/A, not applicable; NDI, neck disability index; NR, not reported; PCF, posterior cervical fusion; RoB = risk of bias; SD, standard deviation; VAS, visual analogue scale.
[a]As reported by authors.
[b]NDI and VAS scores are estimated from graphs for Peppers 2017 and Vanichkachorn 2016.
[c]NDI and VAS were reported on a 0-100 scale with higher scores indicating worse disability and pain.
[d]Authors do not clearly report on numbers of subjects at baseline and 24 months; n's of 182 and 135 respectively are assumed for this report.
[e]Authors reported 24-month fusion results in 2 different ways. We reported "24-month minimum results" that included patients who fused at 12 months and did not receive radiographic follow-up at 24 months. Results for those received 24-month radiographic follow-up only were 82% (138/169).

Discussion

Achieving solid arthrodesis is an important goal of spine fusion surgery when there are structural instability or compromise from degenerative spine disease, spinal deformity, spine trauma, infections, or tumors. However, it is also one of the major challenges faced by spine surgeons in complex spine surgery. Pseudoarthrosis can lead to poor clinical outcomes, instrumentation failure, or revision surgery.[1-8] As a result, surgeons and spine researchers are in constant search for the best biological solution or material to help achieve solid arthrodesis

Table 6. Summary Retrospective Comparative Study (McAnany, 2016): Allogenic Cells for Cervical Fusion.

McAnany 2016 Moderately High RoB, Outcome Measure	Osteocel (n = 57), % (n/N)	Vertigraft (n = 57), % (n/N)	RR (95% CI)[a]	P[a]
Fusion (12 months)				
Rate of fusion (all)	87.7 (50/57)	94.7 (54/57)	0.93 (0.83-1.17)	.190
One level	86.2 (25/29)	96.6 (28/29)	0.89 (0.78-1.30)	.160
Two levels	89.3 (25/28)	89.3 (25/28)	1 (0.83-1.41)	1.000
Safety (12 months)				
Nonunion	12.3 (7/57)	5.3 (3/57)	2.33 (0.81-2.30)	NR
Failed fusion requiring revision surgery for symptomatic pseudarthrosis	57 (4/7)[b]	100 (3/3)	0.57 (0.65-1.76)	NR
Treatment-related postoperative complications	0 (0/57)	0 (0/57)	IC	NR

Abbreviations: CI, confidence interval; IC, incalculable; RoB, risk of bias; RR, risk ratio.
[a]Risk ratios calculated by Aggregate Analytics, Inc. (AAI). *P* values are reported by authors.
[b]The other 3 failed fusion patients in the Osteocel group were asymptomatic and followed clinically but did not receive revision surgery.

in spine fusion cases. While ICBG has been long accepted as the "gold standard" graft material for spine fusion as it possesses osteogenic, osteoinductive, and osteoconductive properties that are critical for spine fusions,[13-18] the use of ICBG is limited by concerns of harvest morbidity, cost, and lack of efficacy in elderly or osteoporotic patients.[14,15,19-21] A variety of bone graft substitutes are available for utilization in spine surgery by spine surgeons today with variable fusion results.[13,15,32] Mesenchymal stem cells and osteogenic progenitor cells are one of the many bone graft substitutes available in the market.[17,29,30,32] Unlike other bone graft extenders that mainly possesses osteoinductive or osteoconductive properties, stem cell products are touted to contain the osteogenic property that is critical for the ideal bone graft material but elusive to all bone graft substitutes except for ICBG.[17,29,30,32] With that claim, it is not surprising that many are enthusiastic about the use of stem cell products in spine fusion surgery to achieve arthrodesis.

Mesenchymal stem cells and osteogenic progenitor cells can be derived from patient's own bone marrows or adipose tissues. However, the quantity and quality of the autogenic stem cells may be limited by the patients' age and biology similar to ICBG.[15,16,18] On the other hand, allogenic stem cells can be derived from a donor and expanded in cultures then optimized for osteogenic differentiations in a controlled process.[17,29,30,32] This process may circumvent the concerns about quantity and quality of the cells that can implanted in spine fusion surgeries. A number of animal studies have examined the efficacy of stem cells in spine fusions, and they showed variable fusion rates in those studies. However, the preclinical animal studies indicate that stem cells can achieve fusion rate similar to autograft controls. While animal studies indicate efficacy with stem cells in spine fusion to achieve fusion similar to autograft, it's unclear if similar outcomes and effectiveness can be seen in the clinical setting.

In this systematic review, we examined the currently available literature reports and studies to determine if allogenic stem cells indeed can achieve similar fusion results and clinical outcomes compared to autograft and other widely studied bone

graft products available in the market today. In addition, we researched the data on complications and cost consideration associated with allogenic stem cells in spine surgery. Based on our systematic review, we identified only 11 studies on allogenic stem cells in spine fusions that met our study criteria. The majority of the studies were retrospective case series and only two retrospective cohort studies were identified. Overall, the evidence for the efficacy and safety of allogenic cell sources for lumbar and cervical arthrodesis in the clinical setting was very low, primarily due to the high risk of bias and lack of precision of included studies, most of which were case series. No full economic studies were identified and no evidence on the impact of patient or intervention characteristics on effectiveness or safety was available.

In the lumbar spine, findings from the only retrospective comparative study identified suggests that allogenic Map3 versus rhBMP-2[44] are similar with regard to improvement in function and pain; however, complications were more common with rhBMB-2. While case series report improvement in function and pain compared with baseline and fusion occurred in ≥90% of patients across case series of allogenic cell products, comparative studies are needed to truly evaluate the effectiveness.

In the cervical spine, one retrospective cohort study of patients with clinical (radiculopathy or myelopathy) and radiographic evidence of degenerative cervical spine disease comparing Osteocel with cadaveric allograft (Vertigraft)[45] reported somewhat lower fusion rates in the Osteocel group compared with the allograft controls (88% vs 95%), but statistical significance was not achieved and no data on function or pain were reported. Across case series, allogenic stem cell products appeared to be associated with improved pain and function, however in the absence of methodologically sound comparative studies, conclusions regarding effectiveness or safety are problematic.

Our study suggest that clinicians and spine surgeons should maintain our enthusiasm for allogenic stem cells in spine fusion with caution. We currently lack high-level clinical evidence that is derived from well-designed randomized controlled or

Table 7. Evidence Summary: Overall Quality (Strength) of Evidence for Effectiveness and Safety Allogenic Cells for Lumbar Fusion.[a]

Outcome	Studies N (Treatments)	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Publication or Reporting Bias	Overall Quality of Evidence	Findings
Cohort: ALIF using Map compared with rhBMP-2								
Function (mean difference, ODI, 0-100 scale)	Retrospective cohort I (N = 41) Lee 2017	Yes (−1)[†]	Unknown§	No	No	Undetected	Very low	• No difference in functional improvement between Map3 (24.8) and rhBMP-2 (25.7) at 12 months
VAS pain (mean difference, 0-10 scale)		Yes (−1)[†]	Unknown§	No	No	Undetected	Very low	• No difference in pain improvement between Map3 (2.2) and rhBMP-2 (2.6) between treatments at 12 months
Fusion (%)[b]		Yes (−1)[†]	Unknown§	No⁎*	No	Undetected	Very low	• Fusion was similar between treatment groups at 12 months (90% vs 92%)
Harms, adverse events		Yes (−1)[†]	Unknown§	No	Yes (−1)[‡]	Undetected	Very low	• Postoperative complications were less common following ALIF using Map3 (30%) compared with rhBMP-2 (57%); statistical significance was not reached due in part to small sample size. • Postoperative radiculitis was less common with Map3 (10% vs 38%) • No differences between Map3 and rhBMP-2 in proportion of patients with nonunion, posterior infection or epidural hematoma
Case Series: Osteocel Plus, Trinity Evolution								
Function (mean ODI, % improvement from baseline)	I Case series N = 35 (Tohmeh 2012, Osteocel Plus)	Yes (−1)[†]	Unknown§	No	Yes (−1)[‡]	Undetected	Very low	• Compared with baseline scores, patients experienced a 41% improvement at 12 months; (mean 27.1 ± NR) • Firm conclusions regarding effectiveness require evaluation from comparative studies
Pain (mean VAS % improvement from baseline)	I Case series N = 35 (Tohmeh 2012, Osteocel Plus) I Case Series N = 43 (Musante 2015, Trinity Evolution + autograft)	Yes (−1)[†]	Unknown§	No	Yes (−1)[‡]	Undetected	Very low	• Across both stem cell products, patients experienced improvement in VAS back and leg pain up to 12 months compared with baseline. • Firm conclusions regarding effectiveness require evaluation with comparative studies

(continued)

Table 7. (continued)

Outcome	Studies N (Treatments)	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Publication or Reporting Bias	Overall Quality of Evidence	Findings
Fusion (%)[b] (radiograph)	4 Case series N = 158 (Osteocel Plus ± autograft, Trinity Evolution + local autograft)	Yes (−1)†	No	No**	No	Undetected	Very low	• Across stem cell products and surgical approaches, fusion was achieved in ≥90% of patients at 12 to 14 months.
Harms, adverse events	4 Case series N = 158 (Osteocel Plus ± autograft, Trinity Evolution + local autograft)	Yes (−1)†	Unknown§	No	Yes (−1)‡	Undetected	Very low	• Evidence on harms and adverse events was sparse and poorly reported • 8% to 9% of patients experienced nonunion (4 series, N = 158) • Additional surgeries were done 7% (n = 3) of patients in one series of Trinity Evolution (N = 43) • No intraoperative complications were seen in 2 studies (Tohmeh, Musante) and none related to stem cell products were seen in 2 series (Ammerman, Musante) • Studies likely lacked sufficient power to detect rare events and comparative studies are needed to further evaluate safety

Abbreviations: ALIF, anterior lumbar interbody fusion; NR, not reported; ODI, Oswestry Disability Index (1-100 scale, higher score, worse disability); rhBMP-2, recombinant human bone morphogenetic protein–2; VAS, visual analogue scale (0-10 scale, higher is worse).

aReasons for downgrading quality of evidence (general):

*Serious risk of bias: the majority of studies did not meet one or more criteria of a good-quality randomized controlled trial (see Appendix for details).

†Serious risk of bias: the majority of studies did not meet 2 or more criteria of a good-quality cohort or are case series (see Appendix for details).

‡Small sample size/insufficient power and/or significant variation in estimates (eg, wide confidence intervals [CIs], large standard deviations). Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention. If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial." If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) CI and/or small sample size may result in downgrade.

§Unknown consistency; single study or different measures used across studies. Inconsistency: differing estimates of effects across trials. If point estimates across trials are in the same direction, do not vary substantially, or heterogeneity can be explained, results may not be downgraded for inconsistency.

**Indirect, intermediate, or surrogate outcomes may be downgraded.

††Reporting bias suspected: limited assessment or limited reporting of specific outcomes.

bFusion was defined as evidence of bridging across endplates, or bridging from endplates to interspace disc plugs (Lee). For fusion for Tohmeh 2012: 97.5% (39/40) were assessed with fluoroscopy-guided, level-by-level, radiography (FGX) and 2.5% (1/40) with computed tomography.

Table 8. Evidence Summary: Overall Quality (Strength) of Evidence for Effectiveness and Safety of Allogenic Cells for Cervical Fusion.[a]

Outcome	Studies N (Treatments)	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Overall Quality of Evidence	Findings
Cohort: Osteocel vs Vertigraft							
Fusion (%)	McAnany 2016, N = 114 (Osteocel)	Yes (−1)*	Unknown§	No**	No	Very low	• By 12 months, overall fusion rates were somewhat lower for Osteocel versus Vertigraft (88% vs 95%) with a similar pattern seen among those having single-level procedures (86.2% vs 96.6%); statistical significance was not achieved in either instance.
Harms, adverse events		Yes (−1)*	Unknown§	No	No	Very low	• More nonunion for (12.3% vs 5.3%) at 12 months Osteocel versus Vertigraft, but results did not reach statistical significance. • No treatment-related complications were observed; however, sample size may not be sufficient to identify rare events.
Case series: Osteocel, Trinity Evolution, Vivigen							
Function (mean NDI, % improvement from baseline)	4 Case series (N = 274) Osteocel®, (Eastlack 2014) Trinity Evolution (Peppers 2017, Vanichkachorn 2016), Vivigen (Divi 2017)	Yes (−1)†	Unknown§	No	No	Very low	• Across case series and stem cell preparations, authors report improved function at various time frames compared with baseline. • Firm conclusions regarding effectiveness require evaluation from comparative studies
Mean pain (VAS)		Yes (−1)†	Unknown§	No	Yes (−1)‡	Very low	• Across case series and products, authors report improved neck pain at various time frames compared with baseline. • Firm conclusions regarding effectiveness require evaluation from comparative studies.
Fusion		Yes (−1)†	Unknown§	No**	Yes (−1)‡	Very low	• Fusion frequency varied across time frames and products. At 6 months rates in 2 series of Trinity evolution of 66% and 79% respectively and of 100% in the Vivigen series 100%. By 12 months, fusion was seen in 89% and 94% of Trinity evolution recipients. At 24 months, the series of Osteocel reported fusion frequency of 87%. • Firm conclusions regarding effectiveness require evaluation from comparative studies
Harms, adverse events		Yes (−1)†	Unknown§	No		Undetected	• Nonunion rates varied across time frames and products. Two series of Trinity Evolution reported nonunion rates of 34.4% and 21.4%, with lower rates by 12 months (10.6% and 6.5%); one Osteocel series reported 18% at 24 months with 13% at >24 months. None were observed in the Vivigen series.

Table 8. (continued)

Outcome	Studies N (Treatments)	Serious Risk of Bias	Serious Inconsistency	Serious Indirectness	Serious Imprecision	Overall Quality of Evidence	Findings
							• No revisions were reported across 3 series • Studies likely had sufficient statistical power to detect rare adverse events. • Detail of reported events are found in Appendix E

NDI = neck disability index (1-100 scale, higher score, worse disability); VAS, visual analogue scale (0-10 scale, higher is worse).

[a]Reasons for downgrading quality of evidence (general):

*Serious risk of bias: the majority of studies did not meet one or more criteria of a good-quality randomized controlled trial (see Appendix for details).

†Serious risk of bias: the majority of studies did not meet 2 or more criteria of a good-quality cohort or are case series (see Appendix for details).

‡Small sample size/insufficient power and/or significant variation in estimates (eg, wide confidence intervals [CIs], large standard deviations). Imprecise effect estimate for an outcome: small sample size and/or confidence interval includes both negligible effect and appreciable benefit or harm with the intervention. If sample size is likely too small to detect rare outcomes, evidence may be downgraded twice. If the estimate is statistically significant, it is imprecise if the CI ranges from "mild" to "substantial." If the estimate is not statistically significant, it is imprecise if the CI crosses the threshold for "mild/small" effects. Wide (or unknown) confidence interval and/or small sample size may result in downgrade.

§Unknown consistency; single study or different measures used across studies. Inconsistency: differing estimates of effects across trials; if point estimates across trials are in the same direction, do not vary substantially or heterogeneity can be explained, results may not be downgraded for inconsistency.

**Indirect, intermediate or surrogate outcomes may be downgraded.

††Reporting bias suspected: limited assessment or limited reporting of specific outcomes.

prospective cohort studies on the use of allogenic stem cells in spine fusion surgery. With the global spine biologics market was valued at $1.9 billion in 2014 and projected to reach $2.4 billion by 2020, the cost burden from spine biologic product should be a major concern in the spine community. Currently, allogenic stem cell products are heavily marketed for their osteogenic potential and perceived benefits of achieving fusion, and surgeons are increasing their utilization of allogenic stem cell products. As surgeons increase our usage of these allogenic stem cell products for spine fusions, we need to insist on obtaining the supporting data for its outcomes, efficacy, and safety from quality and well-designed clinical study.

Conclusions

The overall quality (strength) of evidence of effectiveness and safety of allogenic cell for lumbar and cervical arthrodesis was very low, meaning that we have very little confidence that the effects seen are reflective of the true effects. Firm conclusions regarding effectiveness or safety of allogenic stem cells for arthrodesis are not possible without methodologically sound studies comparing such therapies to other treatments such as standard autograft. In addition to methodologically sound comparative studies, there is a need for formal economic evaluation of stem cell use for fusion.

References

1. Adogwa O, Parker SL, Shau D, et al. Long-term outcomes of revision fusion for lumbar pseudarthrosis: clinical article. *J Neurosurg Spine*. 2011;15:393-398.
2. Dickson DD, Lenke LG, Bridwell KH, Koester LA. Risk factors for and assessment of symptomatic pseudarthrosis after lumbar pedicle subtraction osteotomy in adult spinal deformity. *Spine (Phila Pa 1976)*. 2014;39:1190-1195.
3. Kim YJ, Bridwell KH, Lenke LG, Cho KJ, Edwards CC 2nd, Rinella AS. Pseudarthrosis in adult spinal deformity following multisegmental instrumentation and arthrodesis. *J Bone Joint Surg Am*. 2006;88:721-728.
4. Oliveira OR, Martins SP, Lima WG, Gomes MM. The use of bone morphogenetic proteins (BMP) and pseudarthrosis, a literature review. *Rev Bras Ortop*. 2017;52:124-140.
5. Scheer JK, Oh T, Smith JS, et al; International Spine Study Group. Development of a validated computer-based preoperative predictive model for pseudarthrosis with 91% accuracy in 336 adult spinal deformity patients. *Neurosurg Focus*. 2018;45:E11.
6. Kim YJ, Bridwell KH, Lenke LG, Rhim S, Cheh G. Pseudarthrosis in long adult spinal deformity instrumentation and fusion to the sacrum: prevalence and risk factor analysis of 144 cases. *Spine (Phila Pa 1976)*. 2006;31:2329-2336.
7. Kim YJ, Bridwell KH, Lenke LG, Rinella AS, Edwards C 2nd. Pseudarthrosis in primary fusions for adult idiopathic scoliosis: incidence, risk factors, and outcome analysis. *Spine (Phila Pa 1976)*. 2005;30:468-474.
8. Klineberg E, Gupta M, McCarthy I, Hostin R. Detection of pseudarthrosis in adult spinal deformity: the use of health-related quality-of-life outcomes to predict pseudarthrosis. *Clin Spine Surg*. 2016;29:318-322.
9. Amankulor NM, Xu R, Iorgulescu JB, et al. The incidence and patterns of hardware failure after separation surgery in patients with spinal metastatic tumors. *Spine J*. 2014;14: 1850-1859.
10. Emery SE, Hughes SS, Junglas WA, Herrington SJ, Pathria MN. The fate of anterior vertebral bone grafts in patients irradiated for neoplasm. *Clin Orthop Relat Res*. 1994;(300):207-212.
11. Safaee M, Oh T, Barbaro NM, et al. Results of spinal fusion after spinal nerve sheath tumor resection. *World Neurosurg*. 2016;90: 6-13.
12. Zhang M, Appelboom G, Ratliff JK, et al. Radiographic rate and clinical impact of pseudarthrosis in spine radiosurgery for metastatic spinal disease. *Cureus*. 2018;10:e3631.
13. Haws BE, Khechen B, Narain AS, et al. Iliac crest bone graft for minimally invasive transforaminal lumbar interbody fusion: a prospective analysis of inpatient pain, narcotics consumption, and costs. *Spine (Phila Pa 1976)*. 2018;43:1307-1312.
14. Khan SN, Cammisa FP Jr, Sandhu HS, Diwan AD, Girardi FP, Lane JM. The biology of bone grafting. *J Am Acad Orthop Surg*. 2005;13:77-86.
15. Tuchman A, Brodke DS, Youssef JA, et al. Iliac crest bone graft versus local autograft or allograft for lumbar spinal fusion: a systematic review. *Global Spine J*. 2016;6:592-606.
16. Risbud MV, Shapiro IM, Guttapalli A, et al. Osteogenic potential of adult human stem cells of the lumbar vertebral body and the iliac crest. *Spine (Phila Pa 1976)*. 2006;31:83-89.
17. Robbins MA, Haudenschild DR, Wegner AM, Klineberg EO. Stem cells in spinal fusion. *Global Spine J*. 2017;7:801-810.
18. Silcox DH 3rd, Daftari T, Boden SD, Schimandle JH, Hutton WC, Whitesides TE Jr. The effect of nicotine on spinal fusion. *Spine (Phila Pa 1976)*. 1995;20:1549-1553.

19. Arrington ED, Smith WJ, Chambers HG, Bucknell AL, Davino NA. Complications of iliac crest bone graft harvesting. *Clin Orthop Relat Res*. 1996;(329):300-309.

20. Robertson PA, Wray AC. Natural history of posterior iliac crest bone graft donation for spinal surgery: a prospective analysis of morbidity. *Spine (Phila Pa 1976)*. 2001;26:1473-1476.

21. St John TA, Vaccaro AR, Sah AP, et al. Physical and monetary costs associated with autogenous bone graft harvesting. *Am J Orthop (Belle Mead NJ)*. 2003;32:18-23.

22. Anderson KD, Guest JD, Dietrich WD, et al. Safety of autologous human schwann cell transplantation in subacute thoracic spinal cord injury. *J Neurotrauma*. 2017;34:2950-2963.

23. Centeno CJ, Al-Sayegh H, Freeman MD, Smith J, Murrell WD, Bubnov R. A multi-center analysis of adverse events among two thousand, three hundred and seventy two adult patients undergoing adult autologous stem cell therapy for orthopaedic conditions. *Int Orthop*. 2016;40:1755-1765.

24. Curtis E, Martin JR, Gabel B, et al. A first-in-human, phase I study of neural stem cell transplantation for chronic spinal cord injury. *Cell Stem Cell*. 2018;22:941-950.e6.

25. Henriksson HB, Brisby H. Development and regeneration potential of the mammalian intervertebral disc. *Cells Tissues Organs*. 2013;197:1-13.

26. Niemansburg SL, van Delden JJ, Oner FC, Dhert WJ, Bredenoord AL. Ethical implications of regenerative medicine in orthopedics: an empirical study with surgeons and scientists in the field. *Spine J*. 2014;14:1029-1035.

27. Sheyn D, Shapiro G, Tawackoli W, et al. PTH induces systemically administered mesenchymal stem cells to migrate to and regenerate spine injuries. *Mol Ther*. 2016;24:318-330.

28. Wang Z, Perez-Terzic CM, Smith J, et al. Efficacy of intervertebral disc regeneration with stem cells—a systematic review and meta-analysis of animal controlled trials. *Gene*. 2015;564:1-8.

29. Eltorai AE, Susai CJ, Daniels AH. Mesenchymal stromal cells in spinal fusion: current and future applications. *J Orthop*. 2017; 14:1-3.

30. Piccirilli M, Delfinis CP, Santoro A, Salvati M. Mesenchymal stem cells in lumbar spine surgery: a single institution experience about red bone marrow and fat tissue derived MSCs. *J Neurosurg Sci*. 2017;61:124-133.

31. Salamanna F, Sartori M, Brodano GB, et al. Mesenchymal stem cells for the treatment of spinal arthrodesis: from preclinical research to clinical scenario. *Stem Cells Int*. 2017;2017: 3537094.

32. Skovrlj B, Guzman JZ, Al Maaieh M, Cho SK, Iatridis JC, Qureshi SA. Cellular bone matrices: viable stem cell-containing bone graft substitutes. *Spine J*. 2014;14:2763-2772.

33. Agency for Healthcare Research and Quality. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2014.

34. Institute of Medicine. *Finding What Works in Health Care: Standards for Systematic Reviews*. Washington, DC: National Academies Press; 2011.

35. Patient-Centered Outcomes Research Institute (PCORI) Methodology Committee. PCORI Methodology Report. http://www. pcori.org/sites/default/files/PCORI-Methodology-Report.pdf. Published January 2019. Accessed February 8, 2019.

36. Cochrane Collaboration. *Cochrane Handbook for Systematic Reviews of Interventions*. London, England: Cochrane Collaboration; 2008.

37. Wright JG, Swiontkowski MF, Heckman JD. Introducing levels of evidence to the journal. *J Bone Joint Surg Am*. 2003;85-A:1-3.

38. Ofman JJ, Sullivan SD, Neumann PJ, et al. Examining the value and quality of health economic analyses: implications of utilizing the QHES. *J Manag Care Pharm*. 2003;9:53-61.

39. GraphPad Prism (for Windows) [computer program]. Version 7. 00. La Jolla, CA: GraphPad Software; 2016.

40. GitHub. A collection of R functions supporting the text book *Modern Epidemiology, Second Edition*, by Kenneth J. Rothman Greenland Sander. https://github.com/epijim/episheet. Accessed October 13, 2017.

41. Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol*. 2011; 64:401-406.

42. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64:383-394.

43. Guyatt GH, Oxman AD, Schünemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the *Journal of Clinical Epidemiology*. *J Clin Epidemiol*. 2011;64:380-382.

44. Lee DD, Kim JY. A comparison of radiographic and clinical outcomes of anterior lumbar interbody fusion performed with either a cellular bone allograft containing multipotent adult progenitor cells or recombinant human bone morphogenetic protein-2. *J Orthop Surg Res*. 2017;12:126.

45. McAnany SJ, Ahn J, Elboghdady IM, et al. Mesenchymal stem cell allograft as a fusion adjunct in one- and two-level anterior cervical discectomy and fusion: a matched cohort analysis. *Spine J*. 2016;16:163-167.

46. Kerr EJ 3rd, Jawahar A, Wooten T, Kay S, Cavanaugh DA, Nunley PD. The use of osteo-conductive stem-cells allograft in lumbar interbody fusion procedures: an alternative to recombinant human bone morphogenetic protein. *J Surg Orthop Adv*. 2011;20:193-197.

47. Tohmeh AG, Watson B, Tohmeh M, Zielinski XJ. Allograft cellular bone matrix in extreme lateral interbody fusion: preliminary radiographic and clinical outcomes. *ScientificWorldJournal*. 2012;2012:263637.

48. Ammerman JM, Libricz J, Ammerman MD. The role of Osteocel Plus as a fusion substrate in minimally invasive instrumented transforaminal lumbar interbody fusion. *Clin Neurol Neurosurg*. 2013;115:991-994.

49. Musante DB, Firtha ME, Atkinson BL, Hahn R, Ryaby JT, Linovitz RJ. Clinical evaluation of an allogeneic bone matrix containing viable osteogenic cells in patients undergoing one- and two-level posterolateral lumbar arthrodesis with decompressive laminectomy. *J Orthop Surg Res*. 2016;11:63.

50. Tally WC, Temple HT, Subhawong TY, Ganey T. Transforaminal lumbar interbody fusion with viable allograft: 75 consecutive cases at 12-month follow-up. *Int J Spine Surg*. 2018;12:76-84.

51. Eastlack RK, Garfin SR, Brown CR, Meyer SC. Osteocel Plus cellular allograft in anterior cervical discectomy and fusion: evaluation of clinical and radiographic outcomes from a prospective multicenter study. *Spine (Phila Pa 1976)*. 2014;39: E1331-E1337.

52. Peppers TA, Bullard DE, Vanichkachorn JS, et al. Prospective clinical and radiographic evaluation of an allogeneic bone matrix containing stem cells (Trinity EvolutionÒ Viable Cellular Bone Matrix) in patients undergoing two-level anterior cervical discectomy and fusion. *J Orthop Surg Res*. 2017;12:67.

53. Vanichkachorn J, Peppers T, Bullard D, Stanley SK, Linovitz RJ, Ryaby JT. A prospective clinical and radiographic 12-month outcome study of patients undergoing single-level anterior cervical discectomy and fusion for symptomatic cervical degenerative disc disease utilizing a novel viable allogeneic, cancellous, bone matrix (Trinity Evolution™) with a comparison to historical controls. *Eur Spine J*. 2016;25:2233-2238.

54. Divi SN, Mikhael MM. Use of allogenic mesenchymal cellular bone matrix in anterior and posterior cervical spinal fusion: a case series of 21 patients. *Asian Spine J*. 2017;11:454-462.