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

Journal Title: Global Spine Journal
Volume: Volume 9, Number 1_suppl
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1177/2192568219833336
Permanent URL: https://pid.emory.edu/ark:/25593/tqv10

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

Copyright information:
© The Author(s) 2019
This is an Open Access work distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Accessed October 8, 2019 10:32 PM EDT
Allogenic Stem Cells in Spinal Fusion: A Systematic Review

Patrick C. Hsieh, MD¹, Zorica Buser, PhD¹, Andrea C. Skelly, PhD², Erika D. Brodt², Darrel Brodke, MD³, Hans-Joerg Meisel, MD, PhD⁴, Jong-Beom Park, MD⁵, S. Tim Yoon, MD, PhD⁶, and Jeffrey C. Wang, MD¹

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

Creative Commons Non Commercial No Derivs CC BY-NC-ND: This article is distributed under the terms of the Creative Commons Attribution-Non Commercial-NoDerivs 4.0 License (http://www.creativecommons.org/licenses/by-nc-nd/4.0/) which permits non-commercial use, reproduction and distribution of the work as published without adaptation or alteration, without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
The use of autologous iliac crest bone graft (ICBG) has been considered the “gold standard” for spine fusion. 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. Moreover, harvest of ICBG is also associated with increased rate of morbidities such as chronic pain at the harvest site, infection, and hematoma. 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. 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. 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. 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. In addition, when mix with other bone graft extenders as carriers or fillers, they can add osteoconductivity to the final stem cell products. 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 autograft or autograft in the thoracolumbar or cervical spine?

**Key Question 3:** Is there evidence that patient factors (eg, age, smoking, comorbidity 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.

**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. 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. ClinicalTrials.gov was searched to identify studies which may have new publications.

**Data Extraction**

In addition to results for primary and secondary outcomes, data abstraction included patient characteristics, demographics, lifestyle factors (eg, smoking), comorbidities (eg, obesity), interventions (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
Economic studies were evaluated according to The Quality of Health Economic Studies (QHES) instrument developed by Ofman et al. 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. 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. 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. 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.
human bone morphogenetic protein–2 (rhBMP-2),

3c a s es e r i e s

multipotent adult progenitor cells (Map3) versus recombinant

Lumbar Arthrodesis.

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),

3 case series using Osteocel Plus (mesenchymal stem cells with demineralized

bone matrix) alone or in combination with local autograft,

one case series of using Trinity Evolution cellular bone allograft

combined with local autograft, and one case series of VIA

Graft cellular bone matrix. Across studies, mean ages were

50 to 64 years and 37% to 53% of participants were male.

Diabetes was a common comorbidity (~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 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 treat-

ment 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 improve-

ment in both at 12 months relative to baseline. Similarly, the

case series of posterolateral fusion (PLF) using Trinity Evolution

with local autograft (N = 43) reported significant improve-

ment in both VAS back pain and leg pain at 12 months

compared with baseline (Table 3).

One small series of Osteoplus (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 out-

comes (Appendix E).

Across 5 case series of various allogenic cell products and

fusion approaches (N = 158), fusion occurred in ≥90% of

patients and levels by 12 to 14 months (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, 2 case series using Osteocel Plus and 1 using Vivigen allograft. 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.

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

improvement from baseline in total Short Form–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 ≥90% of

patients and levels by 12 to 14 months (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, 2 case series using Osteocel Plus and 1 using Vivigen allograft. 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.

A small series (N = 21) using Vivigen reported an 85.9% decrease in NDI from baseline and 81.4% decrease in VAS
<table>
<thead>
<tr>
<th>Cohort Study, RoB</th>
<th>Demographics (Mean or %)</th>
<th>Diagnosis</th>
<th>Intervention Characteristics, Previous Treatment</th>
<th>Treatment Groups</th>
<th>Levels, %</th>
</tr>
</thead>
</table>
| Lee 2017 Moderately high RoB | N = 41, 66 segments | 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:  
- Group 1: Map3 allograft  
- Group 2: rhBMP-2 allograft | 1, 2, or 3 levels (Levels treated): NR |
| N = 41, 66 segments | | | | | |
| Age: 53 years | | | | | |
| Male: 39% | | | | | |
| Smokers: 20% | | | | | |
| Symptom duration: NR | | | | | |
| Case Series, RoB Demographics (Mean or %) | Diagnosis | Intervention Characteristics, Previous Treatment | Graft Materials | Levels, % |
| Posterior approach | | | | |
| Musante 2016 High RoB | N = 43 | 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 |
| N = 43 | | | | | |
| Age: 65 years | | | | | |
| Male: 37% | | | | | |
| Smokers: 9% | | | | | |
| Symptom duration: NR | | | | | |
| Multiple approaches | | | | |
| Kerr 2011 High RoB | N = 52 | 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% | Levels treated (%): NR |
| N = 52 | | | | | |
| Age: 50 years | | | | | |
| Male: 40.4% | | | | | |
| Smokers: 42% | | | | | |
| Symptom duration: NR | | | | | |
| Talmeh 2012 High RoB | N = 78 | 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  
- 1, 2, or 3 levels: NR  
- Levels treated (%):  
  - L1/2: 4%  
  - L2/3: 16%  
  - L3/4: 26%  
  - L4/5: 43%  
  - L5/S1: 10% | Levels treated (%): NR |
| N = 78 | | | | | |
| Age: 52 years | | | | | |
| Male: 58% | | | | | |
| Smokers: 13% | | | | | |
| Symptom duration: NR | | | | | |
| Minimally invasive transforaminal approach | | | | |
| Ammerman 2013 High RoB | N = 23, 26 levels | 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 | |
| N = 23, 26 levels | | | | | |
| Age: 57 years | | | | | |
| Male: 52% | | | | | |
| Smokers: 4% | | | | | |
| Symptom duration: NR | | | | | |
| Tally 2018 High RoB | N = 75, 85 levels | 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 + βTCP + concentrated BMA  
- 1 or 2-Level  
- 1-level: 87%  
- 2-level: 13%  
- Levels treated (%): NR | |
| N = 75, 85 levels | | | | | |
| Age: 58 years | | | | | |
| Male: 53% | | | | | |
| Smokers: 21% | | | | | |
| Symptom duration: 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 morphogenetic protein-2; RoB, risk of bias; PEEK, polyetheretherketone; PLF, posterolateral fusion; TLIF, transforaminal lumbar interbody fusion; XLIF, extreme lateral interbody fusion.
The largest series, using Osteocel (N = 135 patients with 24-month follow-up) reported significant improvement at 24 months 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 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 of 66% and 79%, respectively, and of 100% in the Vivigen series54; 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
following ALIF using Map3 (30%) compared with rhBMP-2\textsuperscript{44} (57%), but statistical significance was not reached due in part to the small number of patients (Table 2). Postoperative radioculitis 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 10\% of patients.\textsuperscript{46-49} Reporting of other complications was sparse. No intraoperative complications were seen in three surgical procedures were done in 7\% of patients in one series\textsuperscript{49} and may not be related to stem cell product use. Pseudarthrosis leading to revision occurred in 4\% (3/75) in one series.\textsuperscript{50} Surgical intervention at an adjacent level (1 patient) was reported in another small series (N = 40).\textsuperscript{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.\textsuperscript{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\%\textsuperscript{52} and 21.4\%,\textsuperscript{53} with lower rates by 12 months (10.6\% and 6.5\%). While the small series of Vivigen reported no nonunions,\textsuperscript{54} the series of Osteocel reported 18\% at 24 months with 13\% at >24 months.\textsuperscript{51} No revision surgeries were reported across 3 series.\textsuperscript{51-53}

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

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Allograft Cell/Intervention</th>
<th>Author, Year</th>
<th>Mean ± SD (N), or % (n/N)</th>
<th>P Value Based on Δ vs Baseline*</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODI\textsuperscript{b}</td>
<td>XLIF, Osteocel Plus</td>
<td>Tohmeh, 2012</td>
<td>45.7 ± NR (n = 35)</td>
<td>N/A</td>
</tr>
<tr>
<td>12 months</td>
<td></td>
<td></td>
<td>27.1 ± NR (n = 35)</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>VAS back pain\textsuperscript{b}</td>
<td>PLF, Trinity Evolution + local autograft</td>
<td>Musante, 2016</td>
<td>6.3 ± NR (n = 43)</td>
<td>N/A</td>
</tr>
<tr>
<td>Baseline</td>
<td>XLIF, Osteocel Plus</td>
<td>Tohmeh, 2012</td>
<td>7.4 ± NR (n = 35)</td>
<td>N/A</td>
</tr>
<tr>
<td>3 months</td>
<td>PLF, Trinity Evolution + local autograft</td>
<td>Musante, 2016</td>
<td>2.0 ± NR (n = 43)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>12 months</td>
<td>PLF, Trinity Evolution + local autograft</td>
<td>Musante, 2016</td>
<td>2.1 ± NR (n = 43)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>XLI, Osteocel Plus</td>
<td>Tohmeh, 2012</td>
<td>3.4 ± NR (n = 35)</td>
<td>&lt;.05</td>
<td></td>
</tr>
<tr>
<td>VAS leg pain\textsuperscript{b}</td>
<td>PLF with Trinity Evolution</td>
<td>Musante, 2016</td>
<td>7.1 ± NR (n = 43)</td>
<td>N/A</td>
</tr>
<tr>
<td>Baseline</td>
<td>XLIF with Osteocel Plus</td>
<td>Tohmeh, 2016</td>
<td>6.8 ± NR (n = 35)</td>
<td>N/A</td>
</tr>
<tr>
<td>3 months</td>
<td>PLF, Trinity Evolution + local autograft</td>
<td>Musante, 2016</td>
<td>1.5 ± NR (n = 43)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>12 months</td>
<td>PLF, Trinity Evolution + local autograft</td>
<td>Musante, 2016</td>
<td>1.8 ± NR (n = 43)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>XLI, Osteocel Plus</td>
<td>Tohmeh, 2012</td>
<td>3.8 ± NR (n = 35)</td>
<td>&lt;.05</td>
<td></td>
</tr>
<tr>
<td>Fusion</td>
<td>MITLIF, Osteocel Plus + autograft (facet)</td>
<td>Ammerman, 2013</td>
<td>Overall: 91% (21/23)</td>
<td>N/A</td>
</tr>
<tr>
<td>12 months</td>
<td>XLIF, Osteocel Plus</td>
<td>Tohmeh, 2012</td>
<td>90% of levels (55/61 levels)</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>PLF, Trinity Evolution cellular bone allograft + local autograft</td>
<td>Musante, 2016</td>
<td>91% (39/43)</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>MITLIF, VIA Graft Cancellous Sponge soaked in concentrated BMA</td>
<td>Tally, 2018</td>
<td>96% (72/75)</td>
<td>N/A</td>
</tr>
<tr>
<td>Median 14 months</td>
<td>Circumferential (67%), ALIF (17%) or TLIF (16%) with Osetocel Plus</td>
<td>Kerr, 2011</td>
<td>92% (48/52)</td>
<td>N/A</td>
</tr>
</tbody>
</table>

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; XLI, extreme lateral interbody fusion.

*As reported by authors.

*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.

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

Included studies were not designed to evaluate 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\textsuperscript{45}; however, firm conclusions are not possible given the potential for high risk of
bias in this study and lack of comparison with another treatment (Appendices E and F).

No full economic studies comparing allogenic stem cell 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.
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 5. Summary of Function, Pain, and Fusion Outcomes From Case Series: Allogenic Cells for Cervical Fusion.

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>F/U (Months)</th>
<th>Approach/Allograft</th>
<th>Author, Year</th>
<th>Mean ± SD (n)</th>
<th>P Value Based on Δ vs Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Functionb</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDIc</td>
<td>Baseline</td>
<td>ACDF, Trinity Evolution</td>
<td>Peppers, 2017</td>
<td>44 ± NR (n = 40)</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>ACDF, Trinity Evolution</td>
<td>Vanichkachorn, 2016</td>
<td>45 ± NR (n = 31)</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACDF, ACCF or PCF; Vivigen</td>
<td>Divi, 2017</td>
<td>40.3 ± NR (n = 21)</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>6 months</td>
<td>ACDF, Trinity Evolution</td>
<td>Peppers, 2017</td>
<td>23 ± NR (n = 35)</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACDF, Trinity Evolution</td>
<td>Vanichkachorn, 2016</td>
<td>21 ± NR (n = 28)</td>
<td>&lt;.05</td>
<td></td>
</tr>
<tr>
<td>12 months</td>
<td>ACDF, Trinity Evolution</td>
<td>Peppers, 2017</td>
<td>19 ± NR (n = 38)</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACDF, Trinity Evolution</td>
<td>Vanichkachorn, 2016</td>
<td>17 ± NR (n = 31)</td>
<td>&lt;.05</td>
<td></td>
</tr>
<tr>
<td>Mean Δ from baseline NDI</td>
<td>24 months</td>
<td>ACDF, Osteocel</td>
<td>Eastlack, 2014d</td>
<td>-21.2 ± NR (n = 135)</td>
<td>&lt;.001a</td>
</tr>
<tr>
<td>Average % Decrease in NDI</td>
<td>12 months</td>
<td>ACDF, ACCF or PCF; Vivigen</td>
<td>Divi, 2017</td>
<td>85.9% (n = 21)</td>
<td>NR</td>
</tr>
<tr>
<td>Painb</td>
<td>VAS Neckc</td>
<td>Baseline</td>
<td>ACDF, Trinity Evolution</td>
<td>Peppers, 2017</td>
<td>59 ± NR (n = 40)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ACDF, Trinity Evolution</td>
<td>Vanichkachorn, 2016</td>
<td>59 ± NR (n = 31)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ACDF, ACCF or PCF; Vivigen</td>
<td>Divi, 2017</td>
<td>83 ± NR (n = 21)</td>
</tr>
<tr>
<td>6 months</td>
<td>ACDF, Trinity Evolution</td>
<td>Peppers, 2017</td>
<td>18 ± NR (n = 35)</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACDF, Trinity Evolution</td>
<td>Vanichkachorn, 2016</td>
<td>14 ± NR (n = 28)</td>
<td>&lt;.05</td>
<td></td>
</tr>
<tr>
<td>12 months</td>
<td>ACDF, Trinity Evolution</td>
<td>Peppers, 2017</td>
<td>17 ± NR (n = 38)</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ACDF, Trinity Evolution</td>
<td>Vanichkachorn, 2016</td>
<td>11 ± NR (n = 31)</td>
<td>&lt;.05</td>
<td></td>
</tr>
<tr>
<td>Mean Δ from baseline VAS Neck</td>
<td>24 months</td>
<td>ACDF, Osteocel</td>
<td>Eastlack, 2014d</td>
<td>-34 ± NR (n = 135)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Average % Decrease in VASc</td>
<td>12 months</td>
<td>ACDF, ACCF or PCF; Vivigen</td>
<td>Divi, 2017</td>
<td>81.4% (n = 21)</td>
<td>NR</td>
</tr>
<tr>
<td>Mean Δ from baseline VAS armc</td>
<td>24 months</td>
<td>ACDF, Osteocel</td>
<td>Eastlack, 2014d</td>
<td>-35 ± NR (n = 135)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>VAS Left/Right armc</td>
<td>Baseline</td>
<td>ACDF, Trinity Evolution</td>
<td>Vanichkachorn, 2016</td>
<td>Left: 44 ± NR (n = 31)</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Right: 36 ± NR (n = 31)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6 months</td>
<td>ACDF, Trinity Evolution</td>
<td>Peppers, 2017</td>
<td>Left: 13 ± NR (n = 28)</td>
<td>Left: &lt;.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Right: 6 ± NR (n = 28)</td>
<td>Right: &lt;.05</td>
<td></td>
</tr>
<tr>
<td>12 months</td>
<td>ACDF, Trinity Evolution</td>
<td>Peppers, 2017</td>
<td>Left: 12 ± NR (n = 31)</td>
<td>Left: &lt;.05</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Right: 4 ± NR (n = 31)</td>
<td>Right: &lt;.05</td>
<td></td>
</tr>
<tr>
<td>Fusion</td>
<td>6 months</td>
<td>ACDF, Trinity Evolution</td>
<td>Peppers, 2017</td>
<td>65.7% (23/35)</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ACDF, Trinity Evolution</td>
<td>Vanichkachorn, 2016</td>
<td>78.6% (22/28)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ACDF, ACCF or PCF; Vivigen</td>
<td>Divi, 2017</td>
<td>100% (21/21)</td>
</tr>
<tr>
<td>12 months</td>
<td>ACDF with Trinity Evolution</td>
<td>Peppers, 2017</td>
<td>89.4% (34/38)</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ACDF with Trinity Evolution</td>
<td>Vanichkachorn, 2016</td>
<td>93.5% (29/31)</td>
</tr>
<tr>
<td>24 months</td>
<td>ACDF with Osteocel</td>
<td>Eastlack, 2014</td>
<td></td>
<td>87% (157/180)</td>
<td>N/A</td>
</tr>
</tbody>
</table>

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.

aAs reported by authors.
bNDI and VAS scores are estimated from graphs for Peppers 2017 and Vanichkachorn 2016.
cNDI and VAS were reported on a 0-100 scale with higher scores indicating worse disability and pain.
dAuthors 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.
eAuthors 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.
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 autologous 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-244 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>
<thead>
<tr>
<th>McAnany 2016</th>
<th>Osteocel (n = 57), % (n/N)</th>
<th>VertiGraft (n = 57), % (n/N)</th>
<th>RR (95% CI)</th>
<th>P&lt;sup&gt;α&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fusion (12 months)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rate of fusion (all)</td>
<td>87.7 (50/57)</td>
<td>94.7 (54/57)</td>
<td>0.93 (0.83-1.17)</td>
<td>.190</td>
</tr>
<tr>
<td>One level</td>
<td>86.2 (25/29)</td>
<td>96.6 (28/29)</td>
<td>0.89 (0.78-1.30)</td>
<td>.160</td>
</tr>
<tr>
<td>Two levels</td>
<td>89.3 (28/29)</td>
<td>89.3 (28/29)</td>
<td>1 (0.83-1.41)</td>
<td>1.000</td>
</tr>
<tr>
<td><strong>Safety (12 months)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonunion</td>
<td>12.3 (7/57)</td>
<td>5.3 (3/57)</td>
<td>2.33 (0.81-2.30)</td>
<td>NR</td>
</tr>
<tr>
<td>Failed fusion requiring revision surgery for symptomatic pseudarthrosis</td>
<td>57 (4/7)</td>
<td>100 (3/3)</td>
<td>0.57 (0.65-1.76)</td>
<td>NR</td>
</tr>
<tr>
<td>Treatment-related postoperative complications</td>
<td>0 (0/57)</td>
<td>0 (0/57)</td>
<td>IC</td>
<td>NR</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; IC, incalculable; RoB, risk of bias; RR, risk ratio.

<sup>α</sup>Risk ratios calculated by Aggregate Analytics, Inc. (AAI). P values are reported by authors.

<sup>1</sup>The other 3 failed fusion patients in the Osteocel group were asymptomatic and followed clinically but did not receive revision surgery.
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Studies N (Treatments)</th>
<th>Serious Risk of Bias</th>
<th>Serious Inconsistency</th>
<th>Serious Indirectness</th>
<th>Serious Imprecision</th>
<th>Publication or Reporting Bias</th>
<th>Overall Quality of Evidence</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohort: ALIF using Map compared with rhBMP-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Function (mean difference, ODI, 0-100 scale)</td>
<td>Retrospective cohort 1 (N = 41) Lee 2017</td>
<td>Yes (-1)†</td>
<td>Unknown§</td>
<td>No</td>
<td>No</td>
<td>Undetected</td>
<td>Very low</td>
<td>• No difference in functional improvement between Map3 (24.8) and rhBMP-2 (25.7) at 12 months</td>
</tr>
<tr>
<td>VAS pain (mean difference, 0-10 scale)</td>
<td></td>
<td>Yes (-1)†</td>
<td>Unknown§</td>
<td>No</td>
<td>No</td>
<td>Undetected</td>
<td>Very low</td>
<td>• No difference in pain improvement between Map3 (2.2) and rhBMP-2 (2.6) between treatments at 12 months</td>
</tr>
<tr>
<td>Fusion (%)b</td>
<td></td>
<td>Yes (-1)†</td>
<td>Unknown§</td>
<td>No**</td>
<td>No Undetected</td>
<td>Very low</td>
<td>• Fusion was similar between treatment groups at 12 months (90% vs 92%)</td>
<td></td>
</tr>
<tr>
<td>Harms, adverse events</td>
<td></td>
<td>Yes (-1)†</td>
<td>Unknown§</td>
<td>No</td>
<td>Yes (-1)‡</td>
<td>Undetected</td>
<td>Very low</td>
<td>• 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.</td>
</tr>
<tr>
<td>Case Series: Osteocel Plus, Trinity Evolution</td>
<td>I Case series N = 35 (Tohmeh 2012, Osteocel Plus)</td>
<td>Yes (-1)†</td>
<td>Unknown§</td>
<td>No</td>
<td>Yes (-1)‡</td>
<td>Undetected</td>
<td>Very low</td>
<td>• Postoperative radiculitis was less common with Map3 (10% vs 38%)</td>
</tr>
<tr>
<td>Function (mean ODI, % improvement from baseline)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• No differences between Map3 and rhBMP-2 in proportion of patients with nonunion, posterior infection or epidural hematoma</td>
</tr>
<tr>
<td>Pain (mean VAS % improvement from baseline)</td>
<td>I Case series N = 43 (Musante 2015, Trinity Evolution + autograft)</td>
<td>Yes (-1)†</td>
<td>Unknown§</td>
<td>No</td>
<td>Yes (-1)‡</td>
<td>Undetected</td>
<td>Very low</td>
<td>• Compared with baseline scores, patients experienced a 41% improvement at 12 months; (mean 27.1 ± NR)</td>
</tr>
<tr>
<td></td>
<td>I Case series N = 35 (Tohmeh 2012, Osteocel Plus)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Firm conclusions regarding effectiveness require evaluation from comparative studies</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Across both stem cell products, patients experienced improvement in VAS back and leg pain up to 12 months compared with baseline.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>• Firm conclusions regarding effectiveness require evaluation with comparative studies</td>
</tr>
</tbody>
</table>
### Table 7. (continued)

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

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).

†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) 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.
| yyIndirect, intermediate, or surrogate outcomes may be downgraded.
| yyReport bias suspected: limited assessment or limited reporting of specific outcomes.
| yyFusion 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.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Studies</th>
<th>N (Treatments)</th>
<th>Serious Risk of Bias</th>
<th>Serious Inconsistency</th>
<th>Serious Indirectness</th>
<th>Serious Imprecision</th>
<th>Overall Quality of Evidence</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohort: Osteocel vs Vertigraft Fusion (%)</td>
<td>McAnany 2016, N = 114 (Osteocel)</td>
<td>Yes (− I)*</td>
<td>Unknown§</td>
<td>No**</td>
<td>No</td>
<td>Very low</td>
<td>• 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.</td>
<td></td>
</tr>
<tr>
<td>Harms, adverse events</td>
<td></td>
<td>Yes (− I)*</td>
<td>Unknown§</td>
<td>No</td>
<td>No</td>
<td>Very low</td>
<td>• More nonunion for (12.3% vs 5.3%) at 12 months Osteocel versus Vertigraft, but results did not reach statistical significance.</td>
<td></td>
</tr>
<tr>
<td>Case series: Osteocel, Trinity Evolution, Vivigen Function (mean NDI, % improvement from baseline)</td>
<td>4 Case series (N = 274)</td>
<td>Yes (− I)†</td>
<td>Unknown§</td>
<td>No</td>
<td>No</td>
<td>Very low</td>
<td>• Across case series and stem cell preparations, authors report improved function at various time frames compared with baseline.</td>
<td></td>
</tr>
<tr>
<td>Mean pain (VAS)</td>
<td>Osteocel® (Eastlack 2014), Trinity Evolution (Peppers 2017, Vanichkachorn 2016), Vivigen (Divi 2017)</td>
<td>Yes (− I)†</td>
<td>Unknown§</td>
<td>No</td>
<td>Yes (− I)‡</td>
<td>Very low</td>
<td>• Across case series and products, authors report improved neck pain at various time frames compared with baseline.</td>
<td></td>
</tr>
<tr>
<td>Fusion</td>
<td>Yes (− I)†</td>
<td>Unknown§</td>
<td>No**</td>
<td>Yes (− I)‡</td>
<td>Very low</td>
<td>• 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%.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Harms, adverse events</td>
<td>Yes (− I)†</td>
<td>Unknown§</td>
<td>No</td>
<td>Undetected</td>
<td>• 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 &gt;24 months. None were observed in the Vivigen series.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
Table 8. (continued)

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Studies N (Treatments)</th>
<th>Serious Risk of Bias</th>
<th>Serious Inconsistency</th>
<th>Serious Indirectness</th>
<th>Serious Imprecision</th>
<th>Overall Quality of Evidence</th>
<th>Findings</th>
</tr>
</thead>
</table>

- 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).

*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 (e.g., 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.
References


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

