Cervical and lumbar spine fusions are often the best solution for various degenerative conditions. With the increase in fusion procedures every year, the demand for ideal grafts is increasing proportionally. Bone healing and the restoration of segment mobility depend on the graft characteristics. Iliac crest bone graft (ICBG) used for spinal fusion is considered the gold standard due to its osteogenic, osteoconductive, and osteoinductive characteristics. However, the potential morbidity and complications associated with harvesting make it a less-than-perfect solution.

With the evolution of biological alternatives to autograft bone grafts, there has been interest in the development of various synthetic biological materials. Ideal syn-

**Synthetic bone graft versus autograft or allograft for spinal fusion: a systematic review**

Zorica Buser, PhD,\(^1\) Darrel S. Brodke, MD,\(^2\) Jim A. Youssef, MD,\(^3\) Hans-Joerg Meisel, MD, PhD,\(^4\) Sue Lynn Myhre, PhD,\(^5\) Robin Hashimoto, PhD,\(^6\) Jong-Beom Park, MD,\(^6\) S. Tim Yoon, MD, PhD,\(^7\) and Jeffrey C. Wang, MD\(^1\)

\(^1\)Department of Orthopaedic Surgery, Keck School of Medicine, University of Southern California, Los Angeles, California; \(^2\)Department of Orthopedics, University of Utah School of Medicine, Salt Lake City, Utah; \(^3\)Spine Colorado, Durango, Colorado; \(^4\)Department of Neurosurgery, Bergmannstrost Hospital, Halle, Germany; \(^5\)Spectrum Research, Inc., Tacoma, Washington; \(^6\)Department of Orthopaedic Surgery, Uijongbu St. Mary’s Hospital, The Catholic University of Korea School of Medicine, Uijongbu, Korea; and \(^7\)Department of Orthopedics, Emory Spine Center, Emory University, Atlanta, Georgia

The purpose of this review was to compare the efficacy and safety of synthetic bone graft substitutes versus autograft or allograft for the treatment of lumbar and cervical spinal degenerative diseases. Multiple major medical reference databases were searched for studies that evaluated spinal fusion using synthetic bone graft substitutes (either alone or with an autograft or allograft) compared with autograft and allograft. Randomized controlled trials (RCT) and cohort studies with more than 10 patients were included. Radiographic fusion, patient-reported outcomes, and functional outcomes were the primary outcomes of interest.

The search yielded 214 citations with 27 studies that met the inclusion criteria. For the patients with lumbar spinal degenerative disease, data from 19 comparative studies were included: 3 RCTs, 12 prospective, and 4 retrospective studies. Hydroxyapatite (HA), HA+collagen, β-tricalcium phosphate (β-TCP), calcium sulfate, or polymethylmethacrylate (PMMA) were used. Overall, there were no differences between the treatment groups in terms of fusion, functional outcomes, or complications, except in 1 study that found higher rates of HA graft absorption.

For the patients with cervical degenerative conditions, data from 8 comparative studies were included: 4 RCTs and 4 cohort studies (1 prospective and 3 retrospective studies). Synthetic grafts included HA, β-TCP/HA, PMMA, and biocompatible osteoconductive polymer (BOP). The PMMA and BOP grafts led to lower fusion rates, and PMMA, HA, and BOP had greater risks of graft fragmentation, settling, and instrumentation problems compared with iliac crest bone graft.

The overall quality of evidence evaluating the potential use and superiority of the synthetic biological materials for lumbar and cervical fusion in this systematic review was low or insufficient, largely due to the high potential for bias and small sample sizes. Thus, definitive conclusions or recommendations regarding the use of these synthetic materials should be made cautiously and within the context of the limitations of the evidence.
therapeutic grafts should have all of the characteristics of ICBG along with low immunogenicity and no transmission of disease. In recent decades, various allografts and synthetic grafts have been developed, including demineralized bone matrix, collagens, calcium phosphates (hydroxyapatite [HA] and β-tricalcium phosphate [β-TCP]), ceramics, calcium sulfates, and biodegradable polymers.19

The use of those bone graft substitutes or extenders has become a large market in the commercial spine arena and accounts for significant costs associated with the surgical procedure, but without clinical evidence to support the widespread use of such products. Understanding how a particular graft will affect bone biology and healing is the first step to obtaining a solid fusion. Surgeons should always consider the scientific evidence regarding the safety and efficacy of allogenic grafts when using them for spinal fusion surgery. There have been numerous clinical studies on various synthetic grafts, their compatibilities, and potential drawbacks. However, a detailed review of the level of evidence, safety, and efficacy is needed.

The purpose of this systematic review was to evaluate the literature and examine the scientific evidence associated with synthetic biological products used for spinal fusion in regard to fusion rates, functional and patient-reported outcomes, and safety. With respect to fusion rates, patient-reported outcomes, and adverse events in adult patients with degenerative spine disorders, we sought to answer the following questions: Are synthetic bone graft substitutes (either alone or with an autograft or allograft) safer and more effective than an autograft or allograft when used in thoracolumbar or cervical spine fusion? Which, if any, characteristics of patients (e.g., age, sex, workers’ compensation) and surgical procedures (e.g., approach, type of cage, graft location) are associated with better or worse outcomes when thoracolumbar or cervical spine fusion is performed with synthetic bone graft substitutes compared with an autograft or allograft?

**Methods**

**Electronic Literature Search**

PubMed/MEDLINE, EMBASE, and the Cochrane Collaboration Library were searched through November 5, 2013. The study inclusion and exclusion criteria are detailed in Supplemental Tables 1 and 2 and Fig. 1. Briefly, we sought to identify comparative studies (e.g., randomized controlled trials [RCTs], cohort studies) of synthetic bone graft substitutes (either alone or with autograft or allograft) versus autograft or allograft in patients with degenerative spine disease who underwent fusion procedures of the thoracolumbar or cervical spine. For the purpose of a complete review, we also included studies that used polymethylmethacrylate (PMMA), even though PMMA is not used widely because of pseudarthrosis and graft migration. We included studies that used a concurrent control group or a consecutive historical control group at the same institution.

We excluded studies that used mixed treatments in all groups, such as HA in addition to a synthetic bone substitute, or studies that employed a growth factor (e.g., bone morphogenetic proteins) in addition to synthetics, as the effects of the additive material could not be separated from the effect of the graft type of interest. We also excluded case series, case reports, studies with fewer than 10 patients in either comparison group, and studies for which radiographic fusion or clinical outcomes were not reported.

**Data Extraction**

From the included articles, the following were extracted: study design, intervention and control treatment details, patient characteristics, inclusion/exclusion criteria, follow-up duration and the rate of follow-up for each treatment group (if reported or calculable), patient diagnosis, and funding sources (Supplemental Tables 3 and 4). We recorded clinical outcomes, including radiographic fusion; patient-reported outcomes such as the Oswestry disability index, visual analog scale (VAS), Japanese Orthopaedic Association score and scale, Roland-Morris score, Short Form-36, and patient satisfaction; functional outcomes; and complications and adverse events (Supplemental Tables 5 and 6). In studies in which the synthetic and autograft treatments were each applied unilaterally in the same patient, we did not record patient-reported or functional outcomes, as each patient received both treatments. We recorded each study’s definition of fusion if available (Supplemental Tables 5 and 6). Outcomes were compared at the final follow-up because the follow-up time was reported inconsistently across the included studies.

**Study Quality and Overall Strength of the Body of Literature**

Each study was evaluated for bias risk by 2 reviewers using the preset criteria from The Journal of Bone & Joint Surgery, American Volume31 for therapeutic studies, which were modified to delineate criteria associated with methodological quality (Supplemental Table 7).31 The strength of the overall body of evidence with respect to each outcome was determined based on the precepts outlined by the Grading of Recommendation Assessment, Development and Evaluation (GRADE) Working Group34 and recommendations made by the Agency for Healthcare Research and Quality (AHRQ).32 Additional qualitative analysis was performed according to AHRQ-required (risk of bias, consistency, directness, precision) and additional domains (dose-response, strength of association, publication bias).28

The initial strength of the overall body of evidence was considered “high” for RCTs and “low” for observational studies. The body of evidence may be downgraded 1 or 2 levels based on the following criteria: 1) risk of bias due to study limitations; 2) inconsistency (heterogeneity) of results; 3) indirectness of evidence; 4) imprecision of the effect estimates (e.g., wide confidence intervals); and 5) publication or reporting bias. In the case of methodologically strong observational studies, the body of evidence may be upgraded 1 or 2 levels based on the following criteria: 1) large magnitude of effect; 2) dose-response gradient; and 3) whether all plausible biases would decrease the magnitude of an apparent effect. The final overall strength of the body of literature expresses our confidence in the
An overall strength of "high" means that we are very confident that the true effect lies close to the estimated effect. A "moderate" rating means that we are moderately confident in the effect estimate: the true effect is likely to be close to the estimated effect, but there is a possibility that it is substantially different. An overall strength of "low" means that our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate. Finally, a rating of "very low" means that we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimated effect. In addition, this rating may be used if there is no evidence or it is not possible to estimate an effect.

Data Analysis

We performed all analyses on an individual study level due to heterogeneity and inconsistency in reporting between studies. For continuous outcome measures (e.g.,
VAS scores), we reported (or calculated if not calculated by the author) the mean change scores. In RCTs, if the 12 differences between groups were statistically significant (p < 0.05) (or approached statistical significance) we calculated the risk ratio and risk difference for reliable outcomes (e.g., death) if a causal association might be assumed. For dichotomous outcomes (e.g., fusion), we reported (or calculated if not calculated by the author) the percentage of patients with that outcome. For statistical significance an unpaired t-test was used.

Results

The search strategy returned 214 potentially relevant citations with 27 selected studies; 19 studies involved fusion of the thoracolumbar and 8 studies involved fusion of the cervical spine (Fig. 1). The exclusion criteria and distribution of the other studies are presented in Fig. 1 and Supplemental Table 2.

Lumbar Spine

HA Versus Autologous Graft

One RCT17 and 5 nonrandomized cohort studies (4 prospective and 1 retrospective)12,13,20,21 met our inclusion criteria (Supplemental Tables 3). Overall, there was quite a bit of methodological variability between studies, and thus no results were pooled. In the RCT done by Kovovessis et al.17 60 patients were randomized to receive posterior lumbar fusion with instrumentation and either HA plus local bone (LB) and bone marrow (BM) (source not reported) bilaterally (n = 20), HA+LB+BM on 1 side and ICBG on the other (n = 20) or ICBG bilaterally (n = 20; Supplemental Tables 3). The cohort studies12,13,20,21 enrolled 24–130 patients, and the HA used included BondroSHA,13,20,21 Chitra-HABg (80% HA and 20% bioactive glass ceramic),1 and coraline HA (Interpore International Inc. and Biomet).12 In the prospective study, HA was placed with an autograft (ICBG, BM, or LB) unilaterally and ICBG contralaterally. The retrospective cohort study compared HA+LB to LB alone or with ICBG, all of which were placed bilaterally (Supplemental Tables 3). The RCT17 and 3 cohort studies3,20,21 reported no differences between the treatments at 12 months (Supplemental Tables 5 and 8). In contrast, Hsu et al. found that HA+LB yielded lower fusion rates than ICBG alone (58% vs 90%; p < 0.05; Supplemental Tables 5 and 8).12 Furthermore, Acharya et al. discontinued their study early due to an unexpectedly high rate of resorption of the Chitra-HA graft (0% fusion rate).1 Several studies found no differences in the functional outcomes3,13,17 or complication rates1,13,17 between the groups (Supplemental Tables 5, 8, and 9, respectively).

HA+Collagen Versus Autologous Graft

One small RCT10 and 3 nonrandomized cohort studies15,18,25 compared HA+collagen (Healos; DePuy Synthes) to ICBG and/or LB. In the RCT29 patients were treated with instrumented posterior lumbar fusion with Healos+BM+LB (n = 12) or an LB+cancellous allograft (n = 16; unreported source; Supplemental Table 3). The cohort studies enrolled 25–100 patients and compared Healos+BM (with or without ICBG) that was applied unilaterally or bilaterally to autograft (ICBG or LB) that was applied unilaterally or bilaterally (Supplemental Table 3). For all the studies, fusion was evaluated at 24 months in an independent and/or blind manner (Supplemental Tables 5 and 8). The RCT and 2 prospective cohort studies15,25 found no difference in the achieved fusion rate between the groups. However, Kunakornwatt et al., in their retrospective study, found that fewer patients achieved fusion with Healos+BM aspirate than with LB alone (30% vs 63%, respectively; p < 0.02).18 At the 24-month follow-up, no differences were found between the groups, as reported by the RCT (e.g., pain, function) and 1 prospective cohort study2 (Low Back Outcome Score “success,” Prolo Economic “success,” patient satisfaction) (Supplemental Tables 5 and 8). With the exception of donor site pain, there were no differences in complication risks between groups (Supplemental Table 9).25 The most severe complication in the Healos group was death from pulmonary vein embolus in 1 patient.25

β-TCP Versus Autologous Graft

One RCT10 and 3 nonrandomized cohort studies16,24,35 met our inclusion criteria (Supplemental Table 3). The RCT by Dai et al.10 compared patients who underwent single-level posterior lumbar interbody fusion and were randomly assigned to β-TCP+LB (n = 32) or ICBG (n = 30) grafting that was placed bilaterally. In the cohort studies, which included between 35 and 61 patients, the β-TCP+autograft and LB were applied on opposite sides of the spine in the same patient. In all 4 studies, there were no differences in the fusion rates between the treatment groups at the follow-up (the follow-up period differed between the studies and ranged between 12 and 36 months; Supplemental Tables 5 and 8). Furthermore, the RCT10 found no differences between the treatment groups during the 3-year follow-up in terms of function or quality of life (Supplemental Table 8). However, Dai et al. reported that 80% of ICBG patients were still experiencing donor site pain at 6 weeks, and that in some cases it extended up to 36 months postsurgery (Supplemental Table 9). Other surgical complications and reoperation rates were similar between groups. At the same time, a nonrandomized study by Yamada et al. found no donor site complications in the group who received synthetic bone grafts (Supplemental Table 9).15

Calcium Sulfate Versus Autologous Graft

Four nonrandomized studies were identified that compared OsteoSet calcium sulfate pellets (Wright Medical Technologies) (+LB or BM) to ICBG or LB alone.2,6,8,26 Study enrollment ranged from 40 to 115 patients (Supplemental Table 3), and all but Chang et al.5 evaluated fusion in a blinded manner. Three studies reported no differences in the fusion rate between groups at 12 and 35 months (Supplemental Tables 5 and 8).2,6,8 In contrast, Niu et al. reported that significantly fewer patients achieved fusion on the side treated with OsteoSet versus ICBG (41% vs 91%, respectively; p = 0.0014).26 Nevertheless, they found no statistical differences between treatment groups in terms of pain or function during the 1-year follow-up (Supplemen-
tal Table 8). In the retrospective study done by Chang et al., the OsteoSet group experienced significantly less surgical blood loss than the ICBG group (492 ± 51 vs 605 ± 132 ml, respectively; p = 0.02); however, there was no record of the transfusion rates. There were no differences between groups in terms of epidural hematoma or infection rates. Donor site pain and chronic numbness occurred in 14% of ICBG patients (Supplemental Table 9).

PMMA Versus Autologous Graft

In a small retrospective cohort study, 34 31 degenerative scoliosis patients were treated with posterior lumbar interbody fusion and PMMA (source not reported) plus LB or ICBG+LB (Supplemental Table 3). The authors reported no difference between groups in terms of functional improvement or complications, with an average follow-up of 3.8 years (range 2–7.2 years; Supplemental Tables 5 and 8). However, bone cement leakage occurred in 14% of PMMA patients and donor-site pain in 24% of ICBG patients (Supplemental Table 9).

HA+Collagen Versus Allograft in the Lumbar Spine

One small lumbar RCT compared Healos (HA+collagen) to a mixture of local autograft and allograft. Because of the use of Healos, this study was evaluated as synthetics versus autograft.

Cervical Spine

HA Versus Autologous Graft

One small RCT23 and 2 small retrospective cohort studies14,15 were identified (Supplemental Table 4). The RCT patients were randomly assigned to undergo anterior cervical discectomy and fusion (ACDF) with coralline-derived HA (ProOsteon 200, Interpore Cross Intl. Inc.; n = 13) or ICBG (n = 16). Two cohort studies treated 40–45 patients with ACDF and either HA (alone or with LB) or autograft (ICBG or LB). The follow-up period for the RCT was 24 months and between 12 and 17 months for cohort studies. There were no differences in the fusion rates for all studies or patient23 or clinical (Prolo scale) outcomes (Supplemental Tables 6 and 10).

Adverse events were reported in the RCT and 1 retrospective cohort study (Supplemental Table 11). The RCT was stopped early due to radiographic failure of the HA grafts; graft fragmentation was seen in 89% of HA grafts compared with 11% of ICBG grafts (risk difference 78%; 95% CI 58%–98%; p < 0.05). The graft-settling rates were also higher in the HA group (50% vs 11%, respectively, p < 0.05). Furthermore, fusion with HA grafts had a higher risk of instrumentation-related problems (50% vs 21%, respectively), with screw migration through the bone being most commonly observed, and more grafts in the HA group were associated with at least 3° loss of sagittal alignment over the fused levels (50% vs 21%, respectively), although statistical significance was not reached for either outcome. In the nonrandomized study, 78% of ICBG patients experienced donor site pain for at least 1 week.

β-TCP/HA Versus Autologous Graft

Cho et al.9 compared ACDF with cages and 2 grafting materials in an RCT; the cage was packed with Triosite (40% β-TCP plus 60% HA) (n = 50) or morcellized ICBG (n = 50; Supplemental Table 4). Flexion/extension radiographs were used to determine fusion, and if radiography was inconclusive then CT scan was used. There was no indication that fusion was evaluated in an independent and/or blinded manner.

Fusion was observed in all patients by 6 months; however, there were statistical differences between the groups in the initial months (in the 1st month, 57% of Triosite recipients had fusion compared with 81% of ICBG patients; Supplemental Tables 5 and 10). The Japanese Orthopaedic Association score recovery rate was similar between treatment groups at a mean of 1.7 years of follow-up (Supplemental Table 10). Donor site complications occurred in 6% of ICBG patients, including wound infection, thigh numbness, and subcutaneous hematoma (Supplemental Table 11).

PMMA Versus Autologous Graft

Two small RCTs compared the use of PMMA with ICBG for ACDF (n = 20 and 24 received synthetics, and n = 22 and 30 received ICBG, respectively; Supplemental Table 4). In both RCTs, the PMMA group had significantly lower fusion rates than the ICBG group (0%–30% vs 86.3%–93.3%; Supplemental Tables 6 and 10) as measured at 67 and 12 months.3 There were no differences between the treatment groups in terms of patient outcomes (“success” according to the Odom criteria, pain) at 67 and 12 months (Supplemental Table 10). However, several complications in both graft groups were reported (revision, laryngeal or ulnar nerve palsy), and the rates were similar between groups (Supplemental Table 11). Donor site pain occurred in 3% to 64% of ICBG patients. In addition, the mean loss of disc height was significantly lower (p < 0.001) in the PMMA groups (0.79–0.96 mm) than in the ICBG groups (1.1–2.26 mm).5,27

Biocompatible Osteoconductive Polymer Versus ICBG

In the prospective study done by Madawi et al., patients were treated with ACDF using either biocompatible osteoconductive polymer (BOP) block implants (n = 65) or ICBG (n = 50) (Supplemental Table 4).22 BOP (D.T.I. Med. Corp., D.T.I.s.a. Zoning Industriel d’Heppignies) was composed of 50% matrix (75% copolymer 1-vinyl 2-pyrrolidone and methyl methacrylate plus 25% calcium gluconate) and 50% polyamide-6 fibers. There were several differences between the BOP and ICBG treatment groups, including sex, use of the Smith-Robinson procedure, and the mean length of follow-up. None of the BOP recipients achieved fusion, compared with 96% of those who received ICBG (Supplemental Tables 6 and 10). The average follow-up times were 17.9 months (BOP) and 16.4 months (ICBG). No differences between the groups were found in terms of the percentage of patients with “success,” as measured using Odom’s criteria (Supplemental Table 10). At the same time, disc space collapse, as indicated by the burial of the BOP blocks, was observed in all BOP recipients by 1 to 2 months of follow-up. Partial graft protrusion occurred in 8% of BOP and 22% of ICBG recipients. Other complications were less common and similar between groups (Supplemental Table 11).
Heterologous Artificial Graft (Surgibone) Versus ICBG

One moderately sized retrospective cohort study of cervical fusion compared heterologous artificial graft (Surgibone; n = 101) to ICBG (n = 149; Supplemental Table 4). Surgibone (Unilab Inc.) was made of calf bone (≥ 80% being HA and the remainder being protein). They reported fusion in 98% of patients in both treatment groups at a mean follow-up of 35 (Surgibone) and 31 months (ICBG) (Supplemental Tables 6 and 10). “Good clinical outcomes,” which was defined as no symptoms, were observed in approximately 50% of the patients in both groups, while “fair clinical outcome” (some complaints with some surgical benefit) was seen in 38% of heterologous graft recipients and 46% of ICBG recipients (Supplemental Table 10). Savolainen et al. reported no differences in surgical-site complications between groups (infection, hematoma, and severe pain). However, severe ICBG donor site pain and iliac hematoma were reported in 14% and 3.4% of patients, respectively (Supplemental Table 11).

Evidence Summary

The overall quality (strength) of the evidence across studies for all outcomes was considered low or insufficient. Supplemental Tables 12 and 13 summarize the overall quality of evidence for primary outcomes and the factors related to the rating, as described in the Methods. For both lumbar and cervical studies, there was a serious risk of bias when assessing fusion rates, outcomes, and complications.

Discussion

In this systematic review, the overall strength of evidence in evaluating the potential use and superiority of synthetic biological materials was low or insufficient, largely due to the high potential for bias and small sample sizes.

Lumbar Spine

When HA or HA+collagen were compared with autograft, we found only 3 prospective cohort studies1,15,21 to be at relatively low risk of bias (Supplemental Table 7). All 3 studies that controlled for possible confounders included follow-up for ≥ 80% of patients, cointerventions were applied equally, and independent/blind assessments were achieved. However, in RCTs17,20 primary potential sources of high bias were lack of randomization sequencing, allocation concealment, and whether the data were analyzed according to the intention-to-treat principle, as well as failure to control for potential baseline differences between the treatment groups. In the nonrandomized studies, factors leading to a high risk of bias were inadequate follow-up information and failure to explore possible confounders and to ensure that the cointerventions were applied equally. The study done by Kim et al.15 was deemed to be at high risk of bias because HA was the material of choice for patients with a history of bone disease, osteoporosis, T-score < −3, or previous ICBG harvest. Furthermore, when the amount of local bone was insufficient, ICBG or HA were used as supplements.

All β-TCP studies were found to be at high risk of bias due to the lack of previously described evidence (Supplemental Table 7). One of the main concerns was the lack of independent or blind assessments, as well as inadequate information on the patient populations and how they were distributed among different groups. Even though the fusion rates were high in the studies of both Moro-Barrero et al.24 and Yamada et al.,39 patient degenerative spinal conditions varied.24,35 There was heterogeneity in the fusion level,24,35 and 46% of patients were smokers.24

All 3 prospective cohort studies that compared calcium sulfate were at relatively low risk of bias, fulfilling all27 or all but one2,8 of the criteria for a high-quality prospective cohort study. The lower fusion rates in the study done by Niu et al.20 could be explained by early graft resorption and the fact that the open system could not contain BM aspirate. On the other hand, Chang et al. reported a 92.3% fusion rate with the OsteoSet and a decrease in VAS scores, as well as improvement in the Oswestry Disability Index scores. However, this retrospective study6 was at high risk of bias (Supplemental Table 7) due to the lack of an independent or blind assessment. In addition, the percentage of patients who were originally considered for enrollment and failure to control for potential differences in the baseline characteristics were other sources of bias.

A study of PMMA scaffolds had a high risk of bias (Supplemental Table 7) and did not demonstrate any superiority over autologous graft.34

Cervical Spine

Overall, synthetic grafts performed similarly or worse than autologous grafts in achieving fusion. All of the studies have demonstrated a high risk of bias due to the method of patient enrollment, randomization, follow-up, and treatment allocation.

Studies using HA illustrated a high risk of bias (Supplemental Table 7). In the study by Chang et al., while clinical “success” was seen in 82% of HA and 61% of ICBG patients, a lack of statistical significance was likely due to the small sample size.7 Furthermore, nonrandomized studies by Chang et al.7 and Kim et al.15 failed to use an independent or blinded assessment of fusion, which may partially explain the differences in fusion frequency between those studies and RCTs.

Similar to the HA studies, the risk of bias in the β-TCP and PMMA studies was considered high (Supplemental Table 7). Cho et al.9 reported 100% fusion rates in both the β-TCP and ICBG groups at 6 months. However, spinal fusion was defined by radiography, and only in the case of uncertainty was MRI or CT used. In addition, the study was biased by the lack of blind or independent assessment. For PMMA, the rates of fusion were lower or 0 compared with the ICBG. However, Bärlocher et al. excluded patients with less than 12 months of follow-up.5

Even though both the BOP22 and Surgibone30 studies enrolled a moderate number of patients (115 and 250, respectively), the risk of bias was considered very high (Supplemental Table 7) due to the lack of randomization, concealment, loss on follow-up, and discrepancies between the treatment groups. In the BOP study, Madawi et al.25 reported similar functional outcomes for the BOP (75%) and ICBG (80%) groups, but at the same time there was no fusion in the BOP group. Furthermore, some of the BOP implants were larger than the disc space, leading to endplate...
protrusions and disc collapse, as well as sclerosis and osteophyte formation. The fact that randomization was surgeon dependent with no blind or independent assessment contributed to the high risk of bias. Similarly, Savolainen et al. reported identical fusion rates (98%) between groups. However, the differences between the groups were substantial. The autograft group had a 25-month-longer follow-up, 20% more patients underwent multilevel fusion, and almost twice as many patients underwent Cloward's procedure.

Conclusions

Based on an analysis of 27 studies, our systematic review shows that the incidence of bias was very high in almost all studies, with no RCT or randomized study having Level I evidence, thus preventing synthetic grafts from being deemed beneficial. Systematic reviews have several limitations, and the results have to be discussed with caution. However, the approach for this literature review, including study inclusion and analysis criteria, enabled us to show with certainty that across both spinal regions, synthetic grafts performed similarly and that autologous bone graft often performed the same or even better.

Future RCTs and prospective studies with a better approach of minimizing the bias are needed. Blinded or independent assessment, patient randomization, treatment allocation, and long-term follow-up should be the base of any study. Furthermore, the effect of the grafting material and procedure on the adjacent level should be evaluated to prevent adjacent segment disease.

Acknowledgments

Study support was provided by AOSpine’s Research Department. Analytic support for this work was provided by Spectrum Research, Inc., with funding from the AOSpine Foundation. AOSpine is a clinical division of the AO Foundation—an independent, medically guided, nonprofit organization. The AOSpine Knowledge Forums are pathology-focused working groups acting on behalf of AOSpine in their domain of scientific expertise. Each forum consists of a steering committee of up to 10 international spine experts who meet on a regular basis to discuss research, assess the best evidence for current practices, and formulate clinical trials to advance spine care worldwide.

References

Disclosures

The authors report the following. Dr. Brodke is a consultant for Amedica, Vallum, and DePuy Synthes; receives royalties from Amedica, DePuy Synthes, and Medtronic; owns stock in Amedica; and receives fellowship support from AOSpine (paid directly to institution/employer). Dr. Buser is a consultant for Xeno Medical. Dr. Hashimoto has a financial relationship with AOSpine. Dr. Meisel is consultant (money paid to institution) for Regenerate Life Sciences GmbH, Zyga, and DiFusion Codon (previously); receives royalties from Medtronic and Fehling Aesculap (previously); and owns stocks (money paid to institution) in Regenerate Life Sciences GmbH and in DiFusion. Dr. Wang receives royalties from Stryker, Osprey, Aesculap, Biomet, Amedica, Seaspine, and Synthes; owns stock in Fizomed and Alphatech; owns private investments in Prometheus Spine, Paradigm Spine, Beneveneu, NexGen, Amedica, Vertiflex, Electrocore, Surgitech, VG Innovations, CoreSpine, Expanding Orthopaedics, Osprey, Bone Biologics, Curative Biosciences, and PearlRiver; serves on the board of directors of the North American Spine Society (receives nonfinancial reimbursement for travel for board meetings, courses, etc.), North American Spine Foundation (nonfinancial), Cervical Spine Research Society (nonfinancial reimbursement for travel for board meetings), AOSpine/AO Foundation (receives both nonfinancial reimbursement and honoraria for board and educational activities), Collaborative Spine Research Foundation (nonfinancial), Spine, JAAOS, The Spine Journal, Journal of Spinal Disorders and Techniques, and Global Spine Journal; and receives fellowship support from the AO Foundation (spine fellowship funding paid directly to institution/employer). Dr. Yoon owns stock in Phygen, Alphatec, and Meditec; receives royalties from Meditech Advisors and Stryker Spine (paid directly to institution/employer); received a grant from AOSpine (paid directly to institution/employer); received research support from Biomet (research support given to AREF) and has received nonfinancial research support from Nuvasive and Medtronic. Dr. Youssef is a consultant for NuVasive, Integra, Seaspine, Amedica, and HealthTrust; owns royalties through NuVasive, Integra, Amedica, and Osprey Medical; owns stock in NuVasive, Amedica, Vertiflex, Beneveneu, Paradigm Spine, Promethean Surgical, ISD, SpineVest, and Providence Medical; receives clinical or research support from Globus Medical, Vertiflex, NuVasive, and Integra; and has other financial relationships with NuVasive, Osprey Biomedical, Amedica, and Seaspine.

Author Contributions

Conception and design: Buser, Brodke, Youssef, Meisel, Park, Yoon, Wang. Acquisition of data: Hashimoto, Analysis and interpretation of data: Hashimoto. Drafting the article: Buser, Hashimoto. Critically revising the article: Buser, Brodke, Youssef, Meisel, Myhre, Park, Yoon, Wang. Reviewed submitted version of manuscript: Buser. Approved the final version of the manuscript on behalf of all authors: Buser. Administrative/technical/material support: Buser. Study supervision: Wang.

Supplemental Information

Online-Only Content

Supplemental material is available with the online version of this article.


Correspondence

Zorica Buser, Department of Orthopaedic Surgery, Keck School of Medicine, University of Southern California, Elaine Stevely Hoffman Medical Research Center, HMR 710, 2011 Zonal Ave., Los Angeles, CA 90033. email: zbuser@usc.edu.