

Iliac Crest Bone Graft versus Local Autograft or Allograft for Lumbar Spinal Fusion: A Systematic Review

Alexander Tuchman¹ Darrel S. Brodke² Jim A. Youssef³ Hans-Jörg Meisel⁴ Joseph R. Dettori⁵
Jong-Beom Park⁶ S. Tim Yoon⁷ Jeffrey C. Wang⁸

¹ Department of Neurological Surgery, Keck School of Medicine, University of Southern California, Los Angeles, California, United States

² Department of Orthopedics, University of Utah, School of Medicine, Salt Lake City, Utah, United States

³ Durango Orthopedic Associates, P.C./Spine Colorado, Durango, Colorado, United States

⁴ Department of Neurosurgery, Bergmannstrost Hospital, Halle, Germany

⁵ Spectrum Research, Inc., Tacoma, Washington, United States

⁶ Department of Orthopaedic Surgery, Uijeongbu St. Mary's Hospital, The Catholic University of Korea School of Medicine, Uijeongbu, Korea

⁷ Department of Orthopedics, Emory Spine Center, Atlanta, Georgia, United States

⁸ Department of Orthopaedic Surgery, Keck School of Medicine, University of Southern California, Los Angeles, California, United States

Address for correspondence Alexander Tuchman, MD, Department of Neurological Surgery, Keck School of Medicine, University of Southern California, 1200 N. State Street, Suite 3300, Los Angeles, CA 90033, United States (e-mail: alexandertuchman@gmail.com).

Global Spine J 2016;6:592–606.

Abstract

Study Design Systematic review.

Objective To compare the effectiveness and safety between iliac crest bone graft (ICBG) and local autologous bone and allograft in the lumbar spine.

Methods A systematic search of multiple major medical reference databases identified studies evaluating spinal fusion in patients with degenerative joint disease using ICBG, local autograft, or allograft in the thoracolumbar spine.

Results Six comparative studies met our inclusion criteria. A “low” strength of the overall body of evidence suggested no difference in fusion percentages in the lumbar spine between local autograft and ICBG. We found no difference in fusion percentages based on low evidence comparing allograft with ICBG autograft. There were no differences in pain or functional results comparing local autograft or allograft with ICBG autograft. Donor site pain and hematoma/seroma occurred more frequently in ICBG autograft group for lumbar fusion procedures. There was low evidence around the estimate of patients with donor site pain following ICBG harvesting, ranging from 16.7 to 20%. With respect to revision, low evidence demonstrated no difference between allograft and ICBG autograft. There was no evidence comparing patients receiving allograft with local autograft for fusion, pain, functional, and safety outcomes.

Conclusion In the lumbar spine, ICBG, local autograft, and allograft have similar effectiveness in terms of fusion rates, pain scores, and functional outcomes. However,

Keywords

- ▶ iliac crest bone graft
- ▶ autograft
- ▶ allograft
- ▶ lumbar fusion

received
October 14, 2015
accepted
November 17, 2015
published online
January 6, 2016

DOI <http://dx.doi.org/10.1055/s-0035-1570749>.
ISSN 2192-5682.

© 2016 Georg Thieme Verlag KG
Stuttgart · New York

License terms



ICBG is associated with an increased risk for donor site-related complications. Significant limitations exist in the available literature when comparing ICBG, local autograft, and allograft for lumbar fusion, and thus ICBG versus other fusion methods necessitates further investigation.

Introduction

Lumbar spinal fusion continues to experience a rapid increase in utilization. In the United States, lumbar fusions increased by 134% from 1993 to 2003,¹ and doubled again between 1998 and 2008.² Modern spine surgeons have many choices available to them for the bone graft, including autograft, allograft, and various synthetic materials. Autologous iliac crest bone graft (ICBG) is still considered the historical gold standard in lumbar spinal fusion. However, ICBG harvesting is associated with morbidity including infection, hematoma, fracture, impaired wound healing, and donor site pain.³⁻⁷

In the hope of avoiding donor site morbidity, other autologous bone has been advocated as a suitable graft material, such as local bone harvested from the laminae and spinous processes during the decompressive maneuvers of a lumbar surgery.^{8,9} In addition, various allograft materials have been used as a substitute for autologous ICBG. However, questions remain as to the comparative effectiveness and safety of other types of autograft and cadaver allograft compared with autologous ICBG. Therefore, we sought to answer the following key questions:

1. Is autologous ICBG safer and more effective than fusion with local autograft in the lumbar spine?
2. Is autologous ICBG safer and more effective than fusion with cadaver allograft in the lumbar spine?
3. Is local autograft safer and more effective than fusion with cadaver allograft in the lumbar spine?

Materials and Methods

Electronic Literature Search

A systematic search of Medline, Embase, and the Cochrane Collaboration Library was performed for literature published through December 2013. The search results were limited to the studies published in the English language with human subjects and abstracts available. The reference lists of key articles from the search as well as applicable systematic reviews were also systematically checked to identify any additional eligible studies. Comparative studies (e.g., randomized controlled trials [RCTs], cohort studies) of patients with degenerative joint disease undergoing fusion procedures with autograft harvested from areas other than the iliac crest or with cadaver allograft, or a mixture of local autograft and cadaver allograft, compared with fusion with autologous ICBG in the thoracolumbar spine were sought. Studies were included if they used a concurrent control group or a consecutive historical control group (at the same institution). Studies with an inconsecutive historical control or a control group at a different institution were excluded (► **Table 1**).

Studies were excluded if they did not report results separately by treatment group or used a mixed graft with excluded treatments such as demineralized bone matrix, because the effects of the additive material could not be separated from the effect of the graft type. Furthermore, studies that included skeletally immature patients (<18 years of age) or patients with a history of tumor or infection in the implantation site, trauma, fracture, or adolescent scoliosis were all excluded. Case series, case reports, or studies that consisted of few than 10 cases for either comparison group were excluded. Animal, cadaver, and biomechanical studies were also excluded.

Data Extraction

The following data was extracted: (1) study design; (2) patient characteristics; (3) interventions; (4) inclusion/exclusion criteria; (5) follow-up duration; (6) the rate of follow-up for each treatment group (if reported or calculable); (7) patient-reported functional and pain outcomes (Oswestry Disability Index [ODI], visual analog scale [VAS], Japanese Orthopaedic Association [JOA] score and scale, Roland-Morris score, Modified MacNab score, or patient satisfaction); (8) various clinical outcomes defined by the investigators; (9) complications or adverse events; (10) fusion rate; (11) time to fusion; (12) definition of fusion (► **Table 2**); (13) area where bone graft was harvested; (14) type of bone used (i.e., cancellous); (15) preparation methods (i.e., morselization); and (16) preservation method (i.e., freeze-dried or frozen). In the absence of patient-reported or clinical outcomes, radiographic or clinician-defined fusion was used to determine success. Fusion percentages were compared at final follow-up because follow-up times were reported inconsistently across the studies. All extracted data was examined for trends and possible pooling.

Study Quality and Overall Strength of Body of Literature

We assessed the risk of bias for each article using criteria set by *The Journal of Bone and Joint Surgery, American Volume* for studies on therapy.^{10,11} After individual article evaluation, the strength of the overall body of evidence with respect to each outcome was determined based on the precepts outlined by the Grades of Recommendation Assessment, Development and Evaluation (GRADE) Working Group and recommendations made by the Agency for Healthcare Research and Quality (AHRQ).¹²⁻¹⁴

The initial strength of the overall body of evidence was considered “high” if the majority of the studies were RCTs and “low” if otherwise. GRADE recommends downgrading one or two levels based on risk of bias, inconsistency of results,

Table 1 Summary of inclusion and exclusion criteria

Study component	Inclusion	Exclusion
Participants	Patients with degenerative joint disease in the thoracolumbar spine	Skeletally immature patients (<18 y of age) History of tumor in the implantation site Trauma Fracture Infection at the implantation site Adolescent scoliosis
Intervention	Spine fusion surgery using ICBG autograft	
Comparators	Fusion procedures using local autograft Fusion procedures using allograft	Spine fusion surgery using fibular bone or a bone graft substitute in addition to autograft or allograft
Outcomes	“Success” Radiographic fusion Patient-reported outcomes Clinician-reported outcomes Adverse events	Nonclinical outcomes (other than radiographic fusion)
Study design	Comparative studies to address efficacy, effectiveness, and safety Well-designed prospective case series designed specifically to evaluate complications will be considered only if no comparative studies are available	Nonclinical studies Case reports Case series (unless there are no comparative studies) Studies with <10 patients per treatment group
Publication	Studies published in English in peer-reviewed journals, published HTAs, or publicly available FDA reports	Abstracts, editorials, letters Duplicate publications of the same study that do not report on different outcomes Single reports from multicenter trials Studies reporting on the technical aspects of biologics use in fusion surgery White papers Narrative reviews Articles identified as preliminary reports when results are published in later versions

Abbreviations: FDA, Food and Drug Administration; HTA, Health Technology Assessment; ICBG, iliac crest bone graft.

Table 2 Characteristics of included lumbar fusion studies by key question

Author (year); study design; LoE	Intervention/control	Characteristics	Inclusion/exclusion criteria	Follow-up (range), n/N (% F/U)	Diagnosis	Funding
Key question 1						
Ohtori et al (2011) ¹⁵ ; RCT; LoE: II	Decompression and PLIF with instrumentation Intervention: • Local bone graft (spinal processes and lamina between L4–L5) (structural) Control: • ICBG autograft (structural)	N = 82 Intervention: • n = 42 • Male = 51% • Age = 66 ± 5.5 y Control: • n = 40 • Male = 49% • Age = 67 ± 6.0 y	Inclusion: • Fusion with bone graft • L4 degenerative spondylolisthesis with spinal stenosis • Low back and neck pain for ≥ 12 mo Exclusion: • Previous back surgery • Spinal tumor, infection, trauma	(2–5 y), % NR	L4 degenerative spondylolisthesis (100%)	None stated
Ito et al (2013) ¹⁸ ; retrospective cohort; LoE: III ^a	PLIF with radiolucent carbon fiber cages Intervention: • Local bone graft (from laminectomy) (Morselized) Control: • ICBG autograft (morselized)	N = 109 Intervention: • n = 56 • Male = 58.9% • Age = 48.6 ± 15.3 y Control: • n = 53 • Male = 49.1% • Age = 50.1 ± 13.4 y	Inclusion: • PLIF performed at one level • Radiolucent carbon fiber cages inserted Exclusion: • Previous back surgery, except recurrent disk herniation • Multilevel procedure • Use of metal cages • Lumbar spine spondylolysis • Patients > 65 y old	Mean 50 mo (24–60 mo), % NR	Disk hernia (11%) Spondylolisthesis (47%) Canal stenosis (42%)	None stated
Sengupta et al (2006) ¹⁹ ; retrospective cohort study; LoE: III	Decompression and PLIF with instrumentation Intervention: • Local bone graft (obtained from decompression) (morselized) Control: • ICBG autograft (unspecified preparation)	N = 76 Intervention: • n = 40 • Male = 38% • Age = 60 Control: • n = 36 • Male = 31% • Age = 60 Total: 2+ levels: • Decompression: 45% • Fusion: 33%	Inclusion: • NR Exclusion: • NR	Mean 28 mo (24–72 mo) 76/109 (68%)	Stenosis (61%) Degenerative spondylolisthesis with stenosis (16%) Isthmic spondylolisthesis with disk degeneration (16%) Degenerative scoliosis (>20 degrees) with stenosis (7%) Previous spine surgery (28%)	None stated
Key question 2						
Gibson et al (2002) ²⁰ ; RCT; LoE: II	PLIF with instrumentation Intervention: • Femoral head allograft (morselized, frozen) Control: • ICBG autograft (morselized)	N = 69 Male = 49.3% Age = 42.4 y (19.5–70.2) Intervention: • n = 37 • Male = NR • Age = 43.88 y Control: • n = 32	Inclusion: • Patients presented with low back pain • Trial of conservative treatment with physiotherapy Exclusion: • NR	Mean 6.12 y, 83.4%	NR	None stated

(Continued)

Table 2 (Continued)

Author (year): study design; LoE	Intervention/control	Characteristics	Inclusion/exclusion criteria	Follow-up (range), n/N (% F/U)	Diagnosis	Funding
Putzier et al (2009) ²¹ ; RCT; LoE: II	PLLF with anterior internal fixation and posterior PEEK cage Intervention: • “Cancellous” allograft (structural, freeze-dried) Control: • ICBG autograft (structural)	<ul style="list-style-type: none"> • Male = NR • Age = 40.68 y Intervention: <ul style="list-style-type: none"> • n = 20^b • Male = 50% • Age = 45.5 y Control: <ul style="list-style-type: none"> • n = 20^b • Male = 55% • Age = 45.4 y 	Inclusion: <ul style="list-style-type: none"> • Persistent lumbosacral and/or pseudoradicular complaints after unsuccessful conservative therapy (≥ 6 mo) • Osteochondrosis in Modic stage ≥ 2 with residual disk height < 7 mm, from idiopathic intervertebral disk degeneration (L4–L5 or L5–S1) Exclusion: <ul style="list-style-type: none"> • Degeneration of adjacent segments or additional degenerative findings • Spinal deformities or destructive processes • Previous lumbar spine surgery • Long-term medication with corticoids or nonsteroidal anti-inflammatory drugs • Chronic pain \geq stage II (Gerbershagen) • Osteoporosis, kidney or liver disease, tumors, BMI > 30 kg/m², pregnant or chronic nicotine, alcohol, or drug abuse 	F/U to 12 mo, 40/44 (90.9%)	Isthmic spondylolisthesis, $<$ grade I (25%)	None stated
Wimmer et al (1999) ²² ; retrospective cohort study; LoE: III	Combined A/P fusion with posterior instrumentation Intervention: <ul style="list-style-type: none"> • Femoral head allograft (morselized, frozen) Control: <ul style="list-style-type: none"> • ICBG autograft (morselized) 	N = 104 Male = 46.8% Intervention: <ul style="list-style-type: none"> • n = 39 • Male = NR • Age = 42 y Control: <ul style="list-style-type: none"> • n = 65 • Male = NR • Age = 40 (9–42 y) 	Inclusion: <ul style="list-style-type: none"> • Patients with painful isthmic spondylolisthesis Exclusion: <ul style="list-style-type: none"> • Previous spinal surgery 	Mean 4 y (3–8 y), % NR	Isthmic spondylolisthesis: <ul style="list-style-type: none"> • Slip of L4–L5 (33%) • Slip of L5–S1 (67%) 	None stated
Key question 3						
No studies identified						

Abbreviations: A/P, anterior/posterior; BMI, body mass index; F/U, follow up; ICBG, iliac crest bone graft; LoE, level of evidence; NR, not reported; PEEK, polyetheretherketone; PLLF, posterior lumbar interbody fusion; PLLF, posterolateral lumbar fusion; RCT, randomized control trial.

^aLarge overlap in patient population with study by Ito et al (2010)⁴⁸; this 2013 article excluded patients > 65 years old, most recent publication included.

^bDemographic results reported for patients available for follow-up only.

indirectness of evidence, imprecision of the effect estimates, or suspected publication bias. Alternatively, the body of evidence could be upgraded one or two levels based on a large magnitude of effect or a dose-response gradient. An overall strength of high means that we are very confident that the true effect lies close to that of the estimated effect. A “moderate” rating means that although the true effect is likely to be close to the estimated effect, there is a possibility that it is substantially different. An overall strength of low means that our confidence in the effect estimate is limited, and the true effect may be substantially different from the estimate. Finally, a rating of “insufficient” means that we have very little confidence in the effect estimate; the true effect is likely to be substantially different than the estimated effect. In addition, this rating was used if there was no evidence or it was not possible to estimate an effect.

Data Analysis

The data is summarized in tables and further stratified based on the graft tissue preparation and preservation method. When present, we report mean differences (MDs) of continuous variables and their variance that compare baseline with follow-up values. For studies that did not report MDs, we calculated the MDs when the data was present. Risk proportions (percents) are reported for dichotomous variables by tallying risks as the proportion of patients experiencing an event. When the complication risk was greater in one treatment group compared with another, we calculated the risk ratio and 95% confidence interval using STATA 9.0 (StataCorp, College Station, Texas, United States).

Results

Study Selection

The search strategy yielded 220 potentially relevant citations. Of these, 195 were excluded based on title and/or abstract. Twenty-five were selected for full text review. An additional 19 were excluded based on full text review because they were not a comparison of interest ($n = 10$) or not a population of interest ($n = 4$). Among the six included studies, three compared autologous ICBG with local autograft (key question 1), three compared autologous ICBG with cadaver allograft (key question 2), and none compared local autograft with cadaver allograft (key question 3).

Evidence Available

One RCT by Ohtori et al and two retrospective cohort studies were identified for key question 1 (► **Table 2**).¹⁵⁻¹⁹ All three of these studies involved fusion procedures of the lumbar spine. The RCT ($n = 82$; 50% male; mean age, 66.5 years; main diagnosis, degenerative spondylolisthesis; follow-up, 2 to 5 years) did not specify the preparation method.¹⁵ The other two compared local bone graft with autologous ICBG ($n = 185$; percent male, not reported; mean age, 53.7 years; follow-up, 2 to 6 years). One of the retrospective cohorts used morselized local bone graft as well as morselized ICBG.¹⁸ Sengupta et al used morselized local bone with an unspecified prepara-

tion method to compare with ICBG for posterolateral fusion.¹⁹

Three studies were identified for key question 2,²⁰⁻²² two RCTs and one retrospective cohort.²⁰⁻²² Of these, two used frozen, morselized, femoral-type bone grafts,^{20,22} and the other study used freeze-dried, structural allograft from an unspecified donor location to compare with autologous ICBG.²¹

There were no lumbar studies identified for key question 3.

Efficacy/Effectiveness

Key Question 1

Key question 1 regarded efficacy and effectiveness of local versus autologous ICBG. Three studies reported fusion outcomes for local autograft compared with autologous ICBG.^{15,18,19} In the only RCT, there was no difference in the proportion of patients who obtained fusion between groups receiving local bone or ICBG (83 and 85%; ► **Table 3**).¹⁵ Likewise, in the two retrospective cohort studies, there were no statistical differences comparing fusion between local bone (98 and 96%) and ICBG (65 and 75%; ► **Table 3**).^{18,19}

All three studies compared axial/radicular pain outcomes for local autograft with autologous ICBG.¹⁵⁻¹⁹ In the RCT by Ohtori et al, the MD in pre- and postoperative pain with the VAS (10-point scale) showed no difference between local autograft and autologous ICBG (6.5 and 5.7 for leg and 2.8 and 2.1 for low back, respectively; ► **Table 4**).¹⁵ One retrospective study reported the proportion of patients with reduction in pain of >3 points on the VAS scale for leg pain (75% local bone, 64% ICBG) or low back pain (75% each).¹⁹

There were no significant differences in the patient-reported or clinician-based outcomes (ODI, JOA scale, various clinical ratings) in any of the studies. The RCT by Ohtori et al reported no statistical difference comparing local autograft to autologous ICBG using the ODI with MDs of 22 and 11, respectively.¹⁵ One retrospective cohort reported 87.5% in the local autograft group, and 72.0% in the ICBG group achieved an excellent or good clinical rating ($p = 0.284$).¹⁹ Another retrospective cohort reported no difference in recovery rate of JOA scale (83% local bone, 81% ICBG).¹⁸

Key Question 2

Key question 2 regarded efficacy and effectiveness of allograft versus ICBG. Two studies compared allograft with autologous ICBG. The fusion percentages were similar for studies comparing allograft to autologous ICBG. One RCT ($n = 40$; 52% male; mean age, 45 years, follow-up, 12 months) showed no differences between allograft and autologous ICBG groups (85 and 80% respectively; ► **Table 3**).²¹ One retrospective cohort ($n = 104$; 47% male; age, not reported; follow-up, 4 years) also showed no difference in fusion percentages between groups (92% in the allograft group compared with 95% in the ICBG group).²² There were no apparent differences due to variation in preservation or preparation methods like morselization or freeze-drying/fresh-freezing of allografts.

Pain outcomes were reported in two of the three lumbar fusion studies comparing allograft to autologous ICBG

Table 3 Fusion (%) comparing ICBG with local autograft and ICBG with allograft in the lumbar spine

Author (study design)	Mean age (% male)	F/U mean (range), mo	Fusion definition	Fusion		p Value
				Local autograft or allograft	ICBG	
Local autograft, preparation method unknown						
Ohtori et al (2011) ¹⁵ (RCT)	66.5 y (50%)	(48–60)	Fusion (via radiograph or CT): <1.5 degrees as instability between flexion and extension positions and bridging bone formation across transverse process between adjacent vertebrae	83.3% (35/42)	85% (34/40)	NS
Morselized autograft						
Ito et al (2013) ¹⁸ (retrospective cohort)	49.4 y (54.1%)	50 (24–60)	Fusion (via radiograph): grade 1 or 2 with F/E angle < 5 degrees <ul style="list-style-type: none"> Grade 1: Complete with bone bridge between the upper and lower vertebral bodies Grade 2: Bone bridge not formed, but no translucency around cages with thick fusion mass formation Grade 3: Fusion not achieved with translucency seen around cages Grade 4: Cage sinking into vertebral body or bone resorption around cages 	98% (52/53)	96% (51/53)	NS
Sengupta et al (2006) ¹⁹ (retrospective cohort) ^a	60 y (34.5%)	(24–72)	Fusion (via radiograph): not defined Indeterminate: absence of solid mass, no evidence of halo around implant and absence of motion in F/E Nonunion: not defined	65% (26/40)	75% (27/36)	NS
Structural, freeze-dried allograft						
Putzier et al (2009) ²¹ (RCT)	45.4 y (52.5%)	12	Fusion (via radiograph or CT): positive/negative according to McAfee et al (2001), ³⁷ CT	80% (16/20)	85% (17/20)	NS
Morselized, frozen autograft						
Wimmer et al (1999) ²² (retrospective cohort)	41 y (46.8%)	48 (36–96)	Fusion (via CT): presence of trabecular bony structure spanning site of fusion; no if tethering present (by level)	92% (79/86)	95% (139/146)	NS

Abbreviations: CT, computed tomography; F/E, flexion/extension; F/U, follow-up; ICBG, iliac crest bone graft; NS, not significant; RCT, randomized controlled trial.
^aICBG is structural, local bone is morselized.

Table 4 Pain and clinician-based and patient-reported outcomes following ICBG compared with local autograft in lumbar spinal fusion

Author (study design)	Outcome	Fusion		p Value
		Local autograft or allograft	ICBG	
Local autograft, preparation method unknown				
Ohtori et al (2011) ¹⁵ (RCT)	VAS (0–10) leg MD	6.5 ± 1.6	5.7 ± 1.9	NS
	VAS (0–10) low back MD	2.8 ± 1.0	2.1 ± 1.5	NS
	JOAS (0–3) leg MD	1.6 ± 0.4	1.2 ± 0.4	NS
	JOAS (0–3) low back MD	1 ± 0.5	1.3 ± 0.5	NS
	ODI MD	22 ± 11.4	11 ± 7.2	NS
Morselized autograft				
Ito et al (2013) ¹⁸ (retrospective cohort)	JOAS recovery rate	82.7%	80.5%	NS
Sengupta et al (2006) ¹⁹ (retrospective cohort) ^a	VAS (>3/10) leg pain	75% (30/40)	64% (23/36)	NS
	VAS (>3/10) back pain	75% (31/40)	75% (27/36)	NS
	Excellent/good	87.5% (35/40)	72% (26/36)	NS
	ODI (mean improvement)	36%	32%	
Structural, freeze-dried allograft				
Putzier et al (2009) ²¹ (RCT)	VAS (0–100) MD	42.6 ± 6.3	45.9 ± 7.4	NS
	ODI (excellent/good)	80% (16/20)	75% (15/20)	NS
	Satisfaction	80% (16/20)	75% (15/20)	NS
Morselized, frozen autograft				
Gibson et al (2002) ²⁰ (RCT)	RM back MD	1.9	6.4 ^b	<0.05
	RM leg MD	0.7	1.7	NS
	RM overall MD	0.7	1.4	NS
Wimmer et al (1999) ²² (retrospective cohort)	Improvement	38.9% (15/39)	43.1% (28/65)	NS
	Pain free	33.3% (13/39)	32.3% (21/65)	NS

Abbreviations: ICBG, Iliac crest bone graft; JOAS, Japanese Orthopaedic Association Scale; MD, mean difference; NS, not significant; ODI, Oswestry Disability Index; RCT, randomized controlled trial; VAS, visual analog scale.

^aICBG is structural, local bone is morselized.

^bFive patients who complained of significant donor site pain had much higher back pain than the group average. When their scores were excluded, there was no difference between groups in Roland-Morris (RM) scores at 1-year follow-up.

(► **Table 4**).^{21,22} An RCT by Putzier et al reported no difference between the groups in the change in pain from baseline to follow-up as measured by the VAS (100-point scale; 42.6 in the allograft group and 45.9 for the autograft group).²¹ The retrospective cohort study by Wimmer et al reported no statistical difference in the percentage of patients who had pain improvement (38.5% in the allograft group and 43.1% in the autograft group).²² Pain scores were not influenced by graft preparation or preservation methods.

With respect to patient-reported outcomes in lumbar fusion, two RCTs reported functional results using the ODI, patient satisfaction, and the Roland-Morris score.^{20,21} Gibson et al reported a more favorable functional outcome using the Roland-Morris score for the ICBG group (6.4 overall) compared with the allograft group (1.9 overall). However, five

patients in the autograft group (16.7%) complained of severe donor site pain. When patients with donor site pain at 1 year were excluded from the analysis, this difference between groups disappeared.²⁰ An RCT by Putzier et al, however, reported no significant difference between allograft and autologous ICBG using the ODI: 75 and 80% of patients in each group had excellent or good results. Additionally, the proportion of patients who were satisfied with their surgery was similar between groups (80 and 75%, respectively).²¹ Clinical outcomes were not notably influenced by preparation or preservation of the grafts used.

Key Question 3

Key question 3 regarded efficacy and effectiveness of allograft versus local autograft. No studies were identified evaluating

efficacy and effectiveness of allograft versus local autograft in the lumbar spine.

Safety

Key Question 1

Donor site complications were less frequent in patients receiving local autograft than autologous ICBG (► **Table 5**). Donor site pain at the iliac crest ranged from 6 to 15% in three studies,^{15,18,19} with the RCT by Ohtori et al reporting the highest percentage (15%).¹⁵ There were no donor site complications attributed to local autograft.^{15,18,19} In addition to pain, other donor site complications were reported in the ICBG group: hematoma or seroma percentages ranged from 0 to 5.6%, and sensory loss was 20%.^{15,16,19} No other donor site complications were recorded in those receiving local autograft.

Complications not associated with the donor site were reported inconsistently across studies. There were no differences between treatment groups with respect to infection, dural tears, pedicle screw misplacement, or instrumentation failure (► **Table 5**).

Key Question 2

Donor site complications varied across studies among patients who received autologous ICBG. The risk for donor site pain ranged from 16.7 to 20% at follow-up (► **Table 5**). Other complications observed at the iliac crest donor site included hematoma/seroma (25%).²⁰

There were no reported revisions of lumbar fusion using either allograft or autologous ICBG in the one study reporting this outcome (► **Table 5**).²¹ There were no significant differences between groups with respect to deep infection, superficial wound infection, bleeding from segmental vein, damage to the lateral femoral cutaneous nerve, postsympathetic syndrome, or revision surgery.^{21,22}

Key Question 3

Safety risks for key question 3 were not reported.

Evidence Summary

Key Question 1

Low evidence suggested no difference in fusion percentages between patients receiving local autograft compared with ICBG autograft in the lumbar spine (► **Table 6**).

Low evidence demonstrated no difference in back and leg pain and functional results comparing the two graft types in the lumbar spine. There was moderate evidence that donor site pain and donor site sensory loss occurred more frequently in patients receiving ICBG autograft. There was low evidence that the occurrence of other complications was similar between lumbar local autograft and ICBG autograft.

Key Question 2

Low evidence suggested no difference in fusion percentages between patients receiving allograft compared with ICBG autograft.

There was no difference in pain or function comparing allograft to ICBG autograft in the lumbar spine. The evidence for this conclusion was low.

Donor site pain and hematoma/seroma occurred more frequently in the ICBG autograft group for lumbar fusion procedures. There was low evidence around the estimate (percent) of patients with donor site pain following ICBG harvesting, ranging from 16.7 to 20% (► **Table 6**). With respect to revision, low evidence demonstrated no difference between ICBG autograft and allograft in the lumbar spine.

Key Question 3

There was no evidence comparing patients receiving allograft with local autograft for fusion, pain, or functional and safety outcomes in the lumbar spine.

Discussion

The current clinical literature provides low evidence that ICBG, local autograft, and allograft have similar efficacy in terms of fusion rates and patient outcome measures. Meanwhile, there is moderate to low evidence that harvesting ICBG is associated with a significant morbidity rate. Two modalities were utilized to assess the current literature within a systemic, reproducible, and objective framework. First, individual articles were rated on their class of evidence based on the criteria published by the *Journal of Bone and Joint Surgery, American Volume*.¹⁰ Second, the overall body of evidence with respect to each outcome was stratified according to recommendations of the AHRQ and GRADE working group.¹²⁻¹⁴ This stratification led to predominantly low or insufficient evidence grades regarding the comparative data for ICBG, local autograft, and allograft during lumbar fusion.

Avoiding pseudarthrosis following lumbar fusion procedures is important because long-term fusion is associated with improved patient outcomes.²³ A host of patient- and treatment-associated factors such as smoking status, osteoporosis, diabetes, number of levels treated, use of instrumentation or interbody grafts, and surgical approach are reported to influence fusion rates.²⁴⁻²⁷ This comprehensive literature review focused on the strength of the evidence surrounding the effect choice in bone graft plays on fusion rates and how that is related to patient outcomes and operative morbidity. Controlling for other factors related to fusion was beyond the scope of this systemic review, and the studies included did have significant variation in surgical approaches, use of instrumentation, methodology for assessing fusion, outcome measures reported, indications for surgery, and patient population. Although this interstudy heterogeneity limited our ability to pool the data for a formal meta-analysis, the reviewed studies had sufficient internal control to validate comparison of the differing treatment grafts. Thus, we were able to compare the conclusions of multiple studies and found that they tended to agree that there was no significant difference in terms of fusion rate or outcome when comparing ICBG to local autograft or allograft, and ICBG was associated with a higher incidence of donor site morbidity.

Table 5 Complications at final follow-up^a comparing local autograft versus ICBG, and allograft versus ICBG in lumbar spinal fusion

Outcome	Study	Fusion	
		Local autograft or allograft	ICBG
Local autograft			
Donor site complications			
Donor site pain	Ohtori et al (2011) ¹⁵	0% (0/42)	15% (6/40) ^b
	Ito et al (2013) ¹⁸	0% (0/53)	11% (6/53)
	Sengupta et al (2006)	0% (0/40)	5.6% (2/36)
Hematoma/seroma	Ohtori et al (2011) ¹⁵	0% (0/42)	0% (0/40)
	Sengupta et al (2006) ¹⁹	0% (0/40)	5.6% (2/36)
Sensory loss	Ohtori et al (2011) ¹⁵	0% (0/42)	20% (8/40) ^c
Other complications			
Deep infection	Ohtori et al (2011) ¹⁵	2.3% (1/42)	0% (0/40)
Superficial infection	Ito et al (2013) ¹⁸	0% (0/56)	1.9% (1/53)
Infection (type NS)	Sengupta et al (2006) ¹⁹	5.0% (2/40)	8.3% (3/36)
Deep vein thrombosis	Sengupta et al (2006) ¹⁹	10.0% (4/40)	8.3% (3/36)
Dural tear	Sengupta et al (2006) ¹⁹	12.5% (5/40)	8.3% (3/36)
Hematoma (spinal canal)	Ohtori et al (2011) ¹⁵	0% (0/42)	0% (0/40)
Pedicle screw misplacement	Ito et al (2013) ¹⁸	7.1% (4/56)	5.6% (3/53)
	Sengupta et al (2006) ¹⁹	2.5% (1/40)	8.3% (3/36)
Instrumentation failure	Sengupta et al (2006) ¹⁹	5.0% (2/40)	2.8% (1/36)
Numbness in buttock	Sengupta et al (2006) ¹⁹	0% (0/40)	8.3% (3/36)
Allograft			
Donor site complications			
Pain	Gibson et al(2002) ²⁰	0% (0/37)	16.7% (5/30)
	Wimmer et al (1999) ²²	0% (0/39)	18.5% (12/65)
	Putzier et al (2009) ²¹	0% (0/20)	20% (4/20)
Hematoma/seroma	Gibson et al (2002) ²⁰	0% (0/20)	25% (5/20)
Other complications			
Deep infection	Wimmer et al (1999) ²²	0% (0/39)	1.5% (1/65)
Superficial infection	Wimmer et al (1999) ²²	0% (0/39)	1.5% (1/65)
Revision	Putzier et al (2009) ²¹	0% (0/20)	0% (0/20)
Bleeding from segmental vein	Wimmer et al (1999) ²²	5.1% (2/39)	4.6% (3/65)
Damage to lateral FCN	Wimmer et al (1999) ²²	0% (0/39)	4.6% (3/65)
Postsympathectomy syndrome	Wimmer et al (1999) ²²	5.1% (2/39)	3.1% (2/65)

Abbreviations: FCN, femoral cutaneous nerve; ICBG, Iliac crest bone graft; NS, not significant.

^aSee ► **Table 2** for final follow-up times.

^b $p = 0.025$.

^c $p = 0.01$.

ICBG has historically been described as the “gold standard” graft material during lumbar fusion, because it possesses the three main factors to encourage new bone growth (osteogenicity, osteoconductivity, and osteoinductivity), it is native to the patient, and it is often used as the control arm for RCTs evaluating lumbar fusion.^{28–30} Although many clinical series have reported high fusion rates,^{24,31–33} there are well-documented disadvantages to using ICBG. Harvesting bone from the iliac crest is associated with increased operative time,

blood loss, and postoperative complaints.^{7,34,35} Thus, other methods to harvest autograft during lumbar fusion have been advocated in recent years.^{15,18,19} For posterior approaches, there is often adequate autograft obtained from the spinous processes, laminae, or facet joints during the associated decompression or approach to the disk space. But using local autograft instead of ICBG has theoretical concerns given the smaller volume and the lower ratio of cancellous to cortical bone, which should result in fewer osteogenic cells and less

Table 6 Summary of the quality of evidence

Outcome	Sample size	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Treatment groups (% or mean difference)		Favors ^a
								Local autograft or allograft	ICBG	
Key question 1: Is autologous ICBG safer and more effective than fusion with local autograft in the thoracolumbar spine?										
Fusion	1 RCT (n = 82)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	83.3%	85%	Neither
	2 retrospective cohorts (n = 182)	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Insufficient due to risk of bias and imprecision	65–98%	75–96%	Neither
Pain VAS leg VAS low back	1 RCT (n = 82)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	6.5	5.7	Neither
	1 RCT (n = 82)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	2.8	2.1	Neither
ODI Mean difference Percent improvement	1 RCT (n = 82)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	22	11	Neither
	1 retrospective cohort (n = 76)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Insufficient due to risk of bias and imprecision	36%	32%	Neither
JOAS recovery rate	1 retrospective cohort (n = 109)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Insufficient due to risk of bias and imprecision	82.7%	80.5%	Neither
Clinical result excellent or good	3 retrospective cohorts (n = 137)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Insufficient due to risk of bias and imprecision	81.3–100%	72–93.3%	Neither
Safety Donor site pain	1 RCT (n = 82)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Moderate due to imprecision	0%	15%	Local autograft
	2 retrospective cohorts (n = 185)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	No serious risk of imprecision	Undetected	Insufficient due to risk of bias	0–7.1%	0–11%	Neither
Donor site hematoma/seroma	1 RCT (n = 82)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Moderate due to imprecision	0%	0%	Neither
	1 retrospective cohort (n = 76)	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Insufficient due to risk of bias and imprecision	0%	5.6%	Neither
Donor site sensory loss	1 RCT (n = 82)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	0%	20%	Local autograft
	1 RCT (n = 82)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	2.3%	0%	Neither
Infection ^e	2 retrospective cohorts (n = 185)	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Insufficient due to imprecision	0–5%	1.9–8.3%	Neither
	1 retrospective cohort (n = 76)	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Insufficient due to imprecision	10%	8.3%	Neither
Deep vein thrombosis	1 retrospective cohort (n = 76)	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Insufficient due to imprecision	12.5%	8.3%	Neither
	1 retrospective cohort (n = 76)	No serious risk of bias	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Insufficient due to imprecision	0%	0%	Neither
Dural tear	1 RCT (n = 82)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	0%	0%	Neither
	1 RCT (n = 82)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	0%	0%	Neither
Hematoma	1 RCT (n = 82)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	0%	0%	Neither
	1 RCT (n = 82)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	0%	0%	Neither
Key question 2: Is autologous ICBG safer and more effective than fusion with cadaver allograft in the thoracolumbar spine?										
Fusion	1 RCT (n = 40)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	80%	85%	Neither

This document was downloaded for personal use only. Unauthorized distribution is strictly prohibited.

Table 6 (Continued)

Outcome	Sample size	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence	Treatment groups (% or mean difference)		Favors ^a
								Local autograft or allograft	ICBG	
Pain VAS Pain free	1 retrospective cohort (n = 104)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Insufficient due to risk of bias and imprecision	92%	95%	
	1 RCT (n = 40)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	42.6	45.9	Neither
ODI excellent or good	1 retrospective cohort (n = 104)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Insufficient due to risk of bias and imprecision	33.3%	32.3%	
	1 RCT (n = 40)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	80%	75%	Neither
Roland-Morris (overall score)	1 RCT (n = 69)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	0.7	1.4	Neither
Safety Donor site pain	2 RCTs (n = 109)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	0%	16.7–20%	Allograft
	1 retrospective cohort (n = 104)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Insufficient due to risk of bias and imprecision	0%	18.5%	
Donor site hematoma/ seroma	1 RCT (n = 69)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	0%	25%	Allograft
Infection ^c	1 retrospective cohort (n = 104)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Insufficient due to risk of bias and imprecision	0%	1.5%	
Revision	1 RCT (n = 40)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Low due to risk of bias and imprecision	0%	0%	Neither
Other ^d	1 retrospective cohort (n = 104)	Serious risk of bias ^b	No serious inconsistency	No serious indirectness	Serious risk of imprecision ^c	Undetected	Insufficient due to risk of bias and imprecision	0–5.1%	3.1%–4.6%	

Abbreviations: ICBG, iliac crest bone graft; JOAS, Japanese Orthopaedic Association Scale; ODI, Oswestry Disability Index; RCT, randomized controlled trial; VAS, visual analog scale.

^aIf evidence was graded insufficient, it was not denoted if one treatment group was favored over the other.

^bSerious risk of bias: the majority of studies did not meet two or more important criteria of a good quality RCT or cohort.

^cSerious imprecision: relatively small sample sizes.

^dOther donor site complications include wound dehiscence, thigh dysesthesia, osteomyelitis, unsightly scarring (most frequent), and donor site morbidity.

^eInfection indicates deep infection, superficial wound infection, or a nonspecified infection.

trabecular area that contribute to the osteogenic and osteoconductive properties of autograft.^{30,36}

Despite these concerns, the current clinical literature supports the assertion that ICBG and local autograft have similar efficacy in achieving lumbar fusion.^{15,18,19} This reported equivalency in fusion rates was seen across differing approaches as well: Sengupta et al used posterolateral fusion¹⁹; Ito et al used a posterior interbody approach¹⁸; and Ohtori et al used an uninstrumented anterior interbody ICBG graft compared with instrumented posterolateral fusion with local autograft.¹⁵ Of note, Sengupta et al did report a statistically significant higher rate of nonunion for multilevel posterolateral fusions that used local autograft compared with ICBG.¹⁹ However, this result was based on data from just 10 patients in the local autograft subgroup and 15 in the ICBG subgroup, so drawing any conclusions from these results alone would be dubious. There were significant limitations noted in the available literature with respect to reporting of fusion rates. Fusion was evaluated at distinctive times and defined differently between studies, with no studies using direct surgical inspection.³⁷ Furthermore, only one study routinely used computed tomography (CT) scans.^{15,38} Multiple clinical series utilizing fine-cut, multiplanar CT scans to access lumbar fusion have demonstrated a strong correlation with findings on open surgical exploration.^{8,9} Static X-rays have been shown to be unreliable at predicting lumbar fusion,^{39–41} and although a lack of motion on dynamic X-rays is highly suggestive of fusion, some degree of motion does not necessarily correlate with pseudarthrosis during surgical exploration.⁴⁰ Also, the one RCT in this group had significant internal variability in surgical approach and use of instrumentation, making it difficult to access the role choice of bone graft played in fusion outcomes.¹⁵

We found no comparative human studies that showed a benefit in using ICBG over allograft despite the purported benefits of the osteoinductive and osteogenic properties of ICBG.^{21,22} Similar results have also been shown in a goat animal model comparing autograft and allograft for lumbar fusion.⁴² Though some reports imply that ICBG results in a shorter time to fusion compared with allograft, this benefit does not seem to have an effect on long-term fusion rates or patient outcome measures.^{21,22,43,44} In fact, many reports have demonstrated that achieving fusion during short-term follow-up does not correlate with patient outcome measures and that nonunion does not become clinically significant until long-term follow-up.^{23,24,31,45} Thus, time to fusion only becomes important if the process is delayed to the point where instrumentation failure or graft subsidence occurs prior to fusion. Allograft does carry a risk of immune-mediated rejection or infection, but directly comparing these risks to a cohort receiving autograft would not be feasible given the extremely low incidence.⁴⁶

Any statement on the relative equivalency of ICBG and allograft for lumbar fusion must be tempered by the limitations of the available literature comparing fusion rates. Only two studies reported fusion rates, one RCT that only included 20 patients in each arm and one retrospective cohort.^{21,22} In fact, all three studies that compared ICBG versus allograft had

serious concerns regarding inadequate sample size to determine differences between the two groups.^{20–22} Although both studies used CT-based criteria to determine fusion and a combined anterior–posterior fusion technique with interbody graft, the RCT only included single-level disease and the retrospective review by Wimmer et al included multilevel fusions.²² Despite multilevel fusion being a well-reported risk factor for nonunion, Wimmer et al reported a higher rate of fusion than the single-level procedures by Putzier et al, highlighting the fact that even when studies use similar interventions and criteria to determine fusion, directly comparing fusion rates between studies may not be valid.^{21,22}

In terms of patient outcome measures and pain scores in general, the reviewed studies found no difference in the effectiveness of ICBG, local autograft, and allograft. Not included in this review are the results from the SPORT trial, which also did not find a difference in outcome measures at 4 years whether ICBG or another graft material was used.⁴⁷ Gibson et al is the only study to report a statistically significant difference in outcome measures.²⁰ The higher Roland-Morris back disability scores in the ICBG group was attributable to pain at the donor site, demonstrating that harvesting ICBG is not without risks. In fact, the ICBG group had a higher rate of donor site complications, including donor site pain, hematoma/seroma, or sensory loss, in all studies that reported these outcomes. Meanwhile, there were no significant differences in non-donor site complications. Again, the data regarding safety information is hampered by the inadequate sample size and >20% loss to follow-up in the majority of the studies.

This systemic review identified significant limitations in the individual studies and overall body of the literature comparing lumbar fusion with ICBG, local autograft, or allograft. Based on the current available literature, there is insufficient evidence to compare the safety and efficacy of lumbar fusion when using local autograft versus allograft (key question 3). In total, three studies met the criteria to compare ICBG to local autograft (key question 1) including one RCT, and three studies (two RCTs) could appropriately compare ICBG to allograft (key question 2). There was limited inter-study consistency in regards to follow-up time, indication, surgical procedure, definition of fusion, and reporting of outcomes and complications, significantly limiting our ability to perform a meta-analysis. Thus, the overall quality of the existing literature comparing ICBG and local autograft or allograft remains limited. Future studies should be aimed at comparing fusion rates, efficacy measures, direct and indirect costs, and safety in a prospective fashion. A power analysis should be part of the initial study design to avoid problems with inadequate sample size seen in the majority of the current studies on this topic. A homogenous study population with regards to surgical indication and procedure performed, along with well-defined outcome measures and consistent definition of fusion, would also be ideal.

Conclusions

Based on low evidence from the current literature, ICBG and local autograft or allograft have similar efficacy in terms of

fusion rates, pain scores, and functional outcomes in the lumbar spine. ICBG use was consistently associated with an increased risk for donor site-related complications across studies. However, there were obvious limitations in the literature, and thereby definitive judgments regarding graft choice must be made carefully within the framework of the current literature. Thus, ICBG versus other fusion methods remains an area of clinical equipoise, and further investigation on this topic with prospective randomized trials is warranted.

Disclosures

Alexander Tuchman: Grant (NuVasive)

Darrel S. Brodke: Royalties (Amedica, Depuy Synthes, Medtronic)

Jim A. Yousef: Consultant (NuVasive, Integra, Amedica, HealthTrust); Royalties (NuVasive, Integra, Amedica, Osprey Medical)

Hans-Joerg Meisel: Consultant (Zyga, Difusion); Royalties (Medtronic, Aesculap, Fehling)

Joseph R. Dettori: Personal fees (AOSpine North America, CSRS)

Jong-Beom Park: none

S. Tim Yoon: Consultant (ISSLS, Stryker, Biomet, Mediatech); Grant (Biomet, NIH, Pfizer); Royalties (Stryker, Biomet, Mediatech); Stock ownership (Alphatech, Phygen, Medyssey, Mediatech)

Jeffrey C. Wang: Royalties (Biomet, Stryker, Alphatech, Synthes, Amedica, Osprey, Aesculap, Seaspine); Personal fees (AO Foundation, NASS, CSRS, CSRF); Stock ownership (Fziomed, Promethean Spine, Paradigm Spine, Benevenue, NexGen, Pioneer, Amedica, Vertiflex, Electrocore, Surgitech, Axiomed, VD Innovations, Corespine, Expanding Orthopaedics, Syndicom, Osprey, Bone Biologics, Curative Biosciences, Peraldiver, Pioneer)

Funding sources

Analytic support for this work was provided by Spectrum Research, Inc. with funding from the AOSpine Foundation.

References

- Cowan JA Jr, Dimick JB, Wainess R, Upchurch GR Jr, Chandler WF, La Marca F. Changes in the utilization of spinal fusion in the United States. *Neurosurgery* 2006;59(1):15–20, discussion 15–20
- Rajae SS, Bae HW, Kanim LE, Delamarter RB. Spinal fusion in the United States: analysis of trends from 1998 to 2008. *Spine (Phila Pa 1976)* 2012;37(1):67–76
- Fernyhough JC, Schimandle JJ, Weigel MC, Edwards CC, Levine AM. Chronic donor site pain complicating bone graft harvesting from the posterior iliac crest for spinal fusion. *Spine (Phila Pa 1976)* 1992;17(12):1474–1480
- Hu R, Hearn T, Yang J. Bone graft harvest site as a determinant of iliac crest strength. *Clin Orthop Relat Res* 1995;(310):252–256
- Kurz LT, Garfin SR, Booth RE Jr. Harvesting autogenous iliac bone grafts. A review of complications and techniques. *Spine (Phila Pa 1976)* 1989;14(12):1324–1331
- Sasso RC, LeHuec JC, Shaffrey C; Spine Interbody Research Group. Iliac crest bone graft donor site pain after anterior lumbar interbody fusion: a prospective patient satisfaction outcome assessment. *J Spinal Disord Tech* 2005;18(Suppl):S77–S81
- Younger EM, Chapman MW. Morbidity at bone graft donor sites. *J Orthop Trauma* 1989;3(3):192–195
- Carreon LY, Djurasovic M, Glassman SD, Sailer P. Diagnostic accuracy and reliability of fine-cut CT scans with reconstructions to determine the status of an instrumented posterolateral fusion with surgical exploration as reference standard. *Spine (Phila Pa 1976)* 2007;32(8):892–895
- Carreon LY, Glassman SD, Schwender JD, Subach BR, Gornet MF, Ohno S. Reliability and accuracy of fine-cut computed tomography scans to determine the status of anterior interbody fusions with metallic cages. *Spine J* 2008;8(6):998–1002
- Wright JG, Swiontkowski MF, Heckman JD. Introducing levels of evidence to the journal. *J Bone Joint Surg Am* 2003;85-A(1):1–3
- Norvell DC, Dettori JR, Fehlings MG, Fourney DR, Chapman JR. Methodology for the systematic reviews on an evidence-based approach for the management of chronic low back pain. *Spine (Phila Pa 1976)* 2011;36(21, Suppl):S10–S18
- Atkins D, Best D, Briss PA, et al; GRADE Working Group. Grading quality of evidence and strength of recommendations. *BMJ* 2004;328(7454):1490
- Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011;64(4):401–406
- West S, King V, Carey TS, et al. Systems to rate the strength of scientific evidence. *Evid Rep Technol Assess (Summ)* 2002;47(47):1–11
- Ohtori S, Koshi T, Yamashita M, et al. Single-level instrumented posterolateral fusion versus non-instrumented anterior interbody fusion for lumbar spondylolisthesis: a prospective study with a 2-year follow-up. *J Orthop Sci* 2011;16(4):352–358
- Arthornthurasook A. Anterior lumbar discectomy: a study of iliac and fibular bone graft. *J Med Assoc Thai* 1992;75(5):273–277
- Grob D, Scheier HJ, Dvorak J. Circumferential fusion of the lumbar and lumbosacral spine. Comparison of two techniques of anterior spinal fusion. *Chir Organi Mov* 1991;76(2):123–131
- Ito Z, Imagama S, Kanemura T, et al. Bone union rate with autologous iliac bone versus local bone graft in posterior lumbar interbody fusion (PLIF): a multicenter study. *Eur Spine J* 2013;22(5):1158–1163
- Sengupta DK, Truumees E, Patel CK, et al. Outcome of local bone versus autogenous iliac crest bone graft in the instrumented posterolateral fusion of the lumbar spine. *Spine (Phila Pa 1976)* 2006;31(9):985–991
- Gibson S, McLeod I, Wardlaw D, Urbaniak S. Allograft versus autograft in instrumented posterolateral lumbar spinal fusion: a randomized control trial. *Spine (Phila Pa 1976)* 2002;27(15):1599–1603
- Putzier M, Strube P, Funk JF, et al. Allogenic versus autologous cancellous bone in lumbar segmental spondylodesis: a randomized prospective study. *Eur Spine J* 2009;18(5):687–695
- Wimmer C, Krismer M, Gluch H, Ogon M, Stöckl B. Autogenic versus allogenic bone grafts in anterior lumbar interbody fusion. *Clin Orthop Relat Res* 1999;(360):122–126
- Kornblum MB, Fischgrund JS, Herkowitz HN, Abraham DA, Berkower DL, Ditkoff JS. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective long-term study comparing fusion and pseudarthrosis. *Spine (Phila Pa 1976)* 2004;29(7):726–733, discussion 733–734
- Fischgrund JS, Mackay M, Herkowitz HN, Brower R, Montgomery DM, Kurz LT. 1997 Volvo Award winner in clinical studies. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective, randomized study comparing decompressive laminectomy and arthrodesis with and without spinal instrumentation. *Spine (Phila Pa 1976)* 1997;22(24):2807–2812

- 25 Fritzell P, Hägg O, Wessberg P, Nordwall A; Swedish Lumbar Spine Study Group. Chronic low back pain and fusion: a comparison of three surgical techniques: a prospective multicenter randomized study from the Swedish lumbar spine study group. *Spine (Phila Pa 1976)* 2002;27(11):1131–1141
- 26 Kurd M, Cohick S, Park A, Ahmadinia K, Lee J, An H. Fusion in degenerative spondylolisthesis: comparison of osteoconductive and osteoinductive bone graft substitutes. *Eur Spine J* 2015;24(5):1066–1073
- 27 Miyazaki M, Tsumura H, Wang JC, Alanay A. An update on bone substitutes for spinal fusion. *Eur Spine J* 2009;18(6):783–799
- 28 Papakostidis C, Kontakis G, Bhandari M, Giannoudis PV. Efficacy of autologous iliac crest bone graft and bone morphogenetic proteins for posterolateral fusion of lumbar spine: a meta-analysis of the results. *Spine (Phila Pa 1976)* 2008;33(19):E680–E692
- 29 Rihn JA, Kirkpatrick K, Albert TJ. Graft options in posterolateral and posterior interbody lumbar fusion. *Spine (Phila Pa 1976)* 2010;35(17):1629–1639
- 30 Vaccaro AR, Chiba K, Heller JG, et al; North American Spine Society for Contemporary Concepts in Spine Care. Bone grafting alternatives in spinal surgery. *Spine J* 2002;2(3):206–215
- 31 Herkowitz HN, Kurz LT. Degenerative lumbar spondylolisthesis with spinal stenosis. A prospective study comparing decompression with decompression and intertransverse process arthrodesis. *J Bone Joint Surg Am* 1991;73(6):802–808
- 32 Thomsen K, Christensen FB, Eiskjaer SP, Hansen ES, Fruensgaard S, Bünger CE. 1997 Volvo Award winner in clinical studies. The effect of pedicle screw instrumentation on functional outcome and fusion rates in posterolateral lumbar spinal fusion: a prospective, randomized clinical study. *Spine (Phila Pa 1976)* 1997;22(24):2813–2822
- 33 Zdeblick TA. A prospective, randomized study of lumbar fusion. Preliminary results. *Spine (Phila Pa 1976)* 1993;18(8):983–991
- 34 Keller EE, Triplett WW. Iliac bone grafting: review of 160 consecutive cases. *J Oral Maxillofac Surg* 1987;45(1):11–14
- 35 Summers BN, Eisenstein SM. Donor site pain from the ilium. A complication of lumbar spine fusion. *J Bone Joint Surg Br* 1989;71(4):677–680
- 36 Campana V, Milano G, Pagano E, et al. Bone substitutes in orthopaedic surgery: from basic science to clinical practice. *J Mater Sci Mater Med* 2014;25(10):2445–2461
- 37 McAfee PC, Boden SD, Brantigan JW, et al. Symposium: a critical discrepancy—a criteria of successful arthrodesis following interbody spinal fusions. *Spine (Phila Pa 1976)* 2001;26(3):320–334
- 38 Choudhri TF, Mummaneni PV, Dhall SS, et al. Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 4: radiographic assessment of fusion status. *J Neurosurg Spine* 2014;21(1):23–30
- 39 Blumenthal SL, Gill K. Can lumbar spine radiographs accurately determine fusion in postoperative patients? Correlation of routine radiographs with a second surgical look at lumbar fusions. *Spine (Phila Pa 1976)* 1993;18(9):1186–1189
- 40 Brodsky AE, Kovalsky ES, Khalil MA. Correlation of radiologic assessment of lumbar spine fusions with surgical exploration. *Spine (Phila Pa 1976)* 1991;16(6, Suppl):S261–S265
- 41 Kant AP, Daum WJ, Dean SM, Uchida T. Evaluation of lumbar spine fusion. Plain radiographs versus direct surgical exploration and observation. *Spine (Phila Pa 1976)* 1995;20(21):2313–2317
- 42 Delawi D, Kruyt MC, Huipin Y, et al. Comparing autograft, allograft, and tricalcium phosphate ceramic in a goat instrumented posterolateral fusion model. *Tissue Eng Part C Methods* 2013;19(11):821–828
- 43 Bishop RC, Moore KA, Hadley MN. Anterior cervical interbody fusion using autogeneic and allogeneic bone graft substrate: a prospective comparative analysis. *J Neurosurg* 1996;85(2):206–210
- 44 Tsuang YH, Yang RS, Chen PQ, Liu TK. Experimental allograft in spinal fusion in dogs. *Taiwan Yi Xue Hui Za Zhi* 1989;88(10):989–994
- 45 Mardjetko SM, Connolly PJ, Shott S. Degenerative lumbar spondylolisthesis. A meta-analysis of literature 1970–1993. *Spine (Phila Pa 1976)* 1994;19(20, Suppl):2256S–2265S
- 46 Mroz TE, Joyce MJ, Steinmetz MP, Lieberman IH, Wang JC. Musculoskeletal allograft risks and recalls in the United States. *J Am Acad Orthop Surg* 2008;16(10):559–565
- 47 Radcliff K, Hwang R, Hilibrand A, et al. The effect of iliac crest autograft on the outcome of fusion in the setting of degenerative spondylolisthesis: a subgroup analysis of the Spine Patient Outcomes Research Trial (SPORT). *J Bone Joint Surg Am* 2012;94(18):1685–1692
- 48 Ito Z, Matsuyama Y, Sakai Y, et al. Bone union rate with autologous iliac bone versus local bone graft in posterior lumbar interbody fusion. *Spine (Phila Pa 1976)* 2010;35(21):E1101–E1105