# Quality and Quantity of Published Studies Evaluating Lumbar Fusion during the Past 10 Years: A Systematic Review

Robert Hart[1] Jeffrey T. Hermsmeyer[2] Rajiv K. Sethi[3] Daniel C. Norvell[2]

[1] Department of Orthopaedics, Oregon Health and Science University, Portland, Oregon, United States
[2] Spectrum Research, Inc., Tacoma, Washington, United States
[3] Department of Orthopaedics and Health Services, Virginia Mason Medical Center, University of Washington, Seattle, Washington, United States

**Address for correspondence** Robert Hart, MD, Department of Orthopaedics, Oregon Health and Science University, Portland, OR 97219, United States (e-mail: hartro@ohsu.edu).

Global Spine J 2015;5:207–218.

## Abstract

**Study Design** Systematic review.

**Clinical Questions** (1) Has the proportion and number of randomized controlled trials (RCTs) as an indicator of quality of evidence regarding lumbar fusion increased over the past 10 years? (2) Is there a difference in the proportion of RCTs among the four primary fusion diagnoses (degenerative disk disease, spondylolisthesis, deformity, and adjacent segment disease) over the past 10 years? (3) Is there a difference in the type and quality of clinical outcomes measures reported among RCTs over time? (4) Is there a difference in the type and quality of adverse events measures reported among RCTs over time? (5) Are there changes in fusion surgical approach and techniques over time by diagnosis over the past 10 years?

**Methods** Electronic databases and reference lists of key articles were searched from January 1, 2004, through December 31, 2013, to identify lumbar fusion RCTs. Fusion studies designed specifically to evaluate recombinant human bone morphogenetic protein-2 or other bone substitutes, revision surgery studies, nonrandomized comparison studies, case reports, case series, and cost-effectiveness studies were excluded.

**Results** Forty-two RCTs between January 1, 2004, and December 31, 2013, met the inclusion criteria and form the basis for this report. There were 35 RCTs identified evaluating patients diagnosed with degenerative disk disease, 4 RCTs evaluating patients diagnosed with degenerative spondylolisthesis, and 3 RCTs evaluating patients with a combination of degenerative disk disease and degenerative spondylolisthesis. No RCTs were identified evaluating patients with deformity or adjacent segment disease.

**Conclusions** This structured review demonstrates that there has been an increase in the available clinical database of RCTs using patient-reported outcomes evaluating the benefit of lumbar spinal fusion for the diagnoses of degenerative disk disease and degenerative spondylolisthesis. Gaps remain in the standardization of reportage of adverse events in such trials, as well as uniformity of surgical approaches used. Finally, continued efforts to develop higher-quality data for other surgical indications for lumbar fusion, most notably in the presence of adult spinal deformity and revision of prior surgical fusions, appear warranted.

**Keywords**
- lumbar spine
- spinal fusion
- evidence-based medicine
- adverse events
- spine surgery

**received**
January 6, 2015
**accepted**
March 23, 2015

**DOI** http://dx.doi.org/10.1055/s-0035-1552984.
**ISSN** 2192-5682.

## Study Rationale and Context

Evidence-based medicine (EBM) emphasizes the prioritization of information from well-designed trials in health care decision making. This term now describes the use of the best clinical evidence as the basis for guidelines for the medical and surgical management of problems on a population level. Well-designed randomized controlled trials (RCTs) are considered the highest-level quality of evidence (level 1) regarding a treatment method. As such, clinicians and payers typically refer to them as justification for performance and coverage of specific treatments.

Lumbar fusion surgery is performed for a variety of spinal pathologies. In addition, lumbar fusion can be achieved via a variety of approaches, including isolated posterior fusion, as well as interbody fusion from posterior, lateral, or anterior approaches.[47] More recently, minimally invasive methods of fusion utilizing all of these approaches have also been devised.[7,31] Despite these improvements in surgical technique, some indications for lumbar fusion surgery, such as in the treatment of axial back pain from degenerative disk disease (DDD), remain controversial.[14,16] Other conditions such as instability, tumor, trauma, or spinal deformity are considered better-proven indications, although there remains significant variability of fusion utilization and technique performed nationally and internationally.[1,14]

Given a relative lack of RCT-quality data, other analyses of billing databases have questioned the indication and benefit of lumbar fusion. However, in many cases these evaluations fail to define the surgical indication and often resort to a relatively nonspecific diagnosis such as "back pain," which leads to increased confusion for health care economists and hospital administrators, many of whom may lack a clinical understanding of surgical diagnoses.[15] Although many surgical patients' complaints may include back pain, a large number are not undergoing surgical fusion exclusively for that symptom but instead are due to associated features such as spinal instability, deformity, or neurologic compression. Thus, large database analyses are not an adequate substitute for higher-quality RCT data.

With the introduction of the Affordable Care Act and increased emphasis on comparative effectiveness research, more attention has been focused on the costs associated with spine care in the United States.[39] Concomitantly, there have been significant technological advances in spinal surgery, increasing the associated costs. Among other issues, questions about the benefits of bone morphogenic protein and incomplete reportage of its complication profile have emerged.[10] It has also recently been shown that reporting of adverse events in cervical total disk trials was inconsistent.[1] All of these features argue for an increase in the quality of clinical research of spine surgical outcomes, both with respect to study design as well as clinical outcome and adverse events recording and reporting.

In this analysis, we set out to determine if there is a difference in the number and proportion of RCTs in the past 10 years among the four most common indications for lumbar spine fusion: DDD, spondylolisthesis, spinal deformity, and adjacent segment disease. We also sought to ascertain whether there has been an improvement in the consistency of clinical outcomes measured among RCTs over time, as well as in the quality of recording and reporting of adverse events. Finally, we also evaluated whether there were consistent changes in fusion surgical approaches reported over the same period.

## Clinical Questions

1. Is the proportion of RCTs as a surrogate for quality of evidence regarding lumbar fusion increasing over the past 10 years?
2. Is there a difference in the proportion of RCTs among the four primary fusion diagnoses (DDD, spondylolisthesis, deformity, and adjacent segment disease) over the past 10 years?
3. Is there a difference in type and quality of clinical outcomes measured among RCTs over time?
4. Is there a difference in type and quality of adverse events measured among RCTs over time?
5. Are there changes in fusion treatment approaches over time by diagnosis over the past 10 years?

## Materials and Methods

**Study design:** Systematic review.
**Search:** PubMed, Cochrane collaboration database, and National Guideline Clearinghouse databases; bibliographies of key articles.
**Dates searched:** January 1, 2004, to December 31, 2013.
**Inclusion criteria:** For clinical questions 1 and 2, a search was done for all study designs and randomized trials separately. For questions 3 to 5, the following criteria were applied: (1) RCTs evaluating lumbar fusion in peer-reviewed journals; (2) patients with any of the following diagnoses undergoing anterior, posterior, circumferential, or transforaminal lumbar fusion: DDD, degenerative spondylolisthesis (DS), adjacent segment disease, or adult spinal deformity; (3) outcomes included patient-reported outcomes, clinician-based outcomes, and adverse events.
**Exclusion criteria:** (1) Fusion studies designed specifically to evaluate recombinant human bone morphogenetic protein-2 or other bone substitutes; (2) revision surgery studies; and (3) nonrandomized comparison studies, case reports, case series, cost-effectiveness studies, prognostic studies for clinical questions 3 to 5.
**Outcomes:** (1) Proportion of RCTs by year and by diagnosis; (2) type of clinical outcomes (i.e., patient-reported versus clinician-based); (3) actual patient-reported outcomes and clinician-based outcomes; (4) type of adverse events; (5) actual adverse events; (6) existence of severity classification for adverse events; and (7) type of fusion approach (i.e., anterior, posterior, circumferential).
**Analysis:** This study was not a comparative effectiveness or safety review; therefore, only descriptive statistics were used to answer the key questions. For clinical questions 1 and 2, the inclusion and exclusion criteria were applied to identify the number of RCTs by year and by diagnosis. This value became

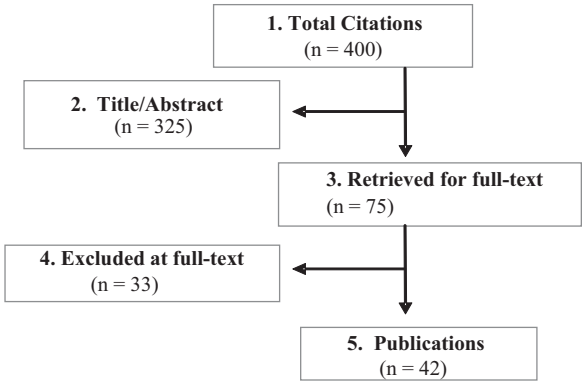


**Fig. 1** Flowchart showing results of literature search.

the numerator of the proportion. The same search was done without the RCT limitation to identify all study designs evaluating fusion. This value became the denominator of the proportion to compute the proportion of RCTs by year and by diagnosis. For the remaining key questions, proportions for each category are reported.

Details about methods can be found in the online supplementary material.

## Results

- We identified 42 RCTs between January 1, 2004, and December 31, 2013, that met the inclusion criteria and form the basis for this report (►**Fig. 1**). See online supplementary material.
- There were 35 RCTs identified evaluating patients diagnosed with DDD, 4 RCTs evaluating patients diagnosed with DS, and 3 RCTs evaluating patients with a combination of DDD and DS (►**Table 1**). No RCTs were identified evaluating patients with deformity or adjacent segment disease.

### Clinical Question 1: Is the Proportion of RCTs as a Surrogate for Quality of Evidence Regarding Lumbar Fusion Increasing over the Past 10 Years?

- The overall proportion of RCTs in the lumbar fusion literature over 10 years was 10.5% ($n$ = 42/400; ►**Fig. 2**).
- The largest proportion of RCTs was in 2004 ($n$ = 5/25; 20%). The next two largest proportions were 2009 ($n$ = 5/31; 16.1%) and 2013 ($n$ = 8/58; 13.8).
- The smallest proportion of RCTs was in 2008 ($n$ = 2/42; 4.8%).
- The other 6 years within the past 10 varied from 8.3 to 9.8%.

### Clinical Question 2: Is There a Difference in the Proportion of RCTs among the Four Primary Fusion Diagnoses (DDD, Spondylolisthesis, Adult Deformity, Adjacent Segment Disease) over the Past 10 Years?

- The overall proportion of RCTs evaluating lumbar fusion in patients with DDD over 10 years was 13.4% ($n$ = 38/284; ►**Fig. 3**).
- The overall proportion of RCTs evaluating lumbar fusion in patients with DS over 10 years was 11.7% ($n$ = 7/60).
- There were no RCTs in the lumbar fusion literature evaluating patients with adult spinal deformity or adjacent segment disease.
- The greatest proportion of fusion RCTs evaluating patients with DDD occurred in the year 2004 ($n$ = 5/21; 23.8%) followed by the year 2013 ($n$ = 7/31; 22.5%).

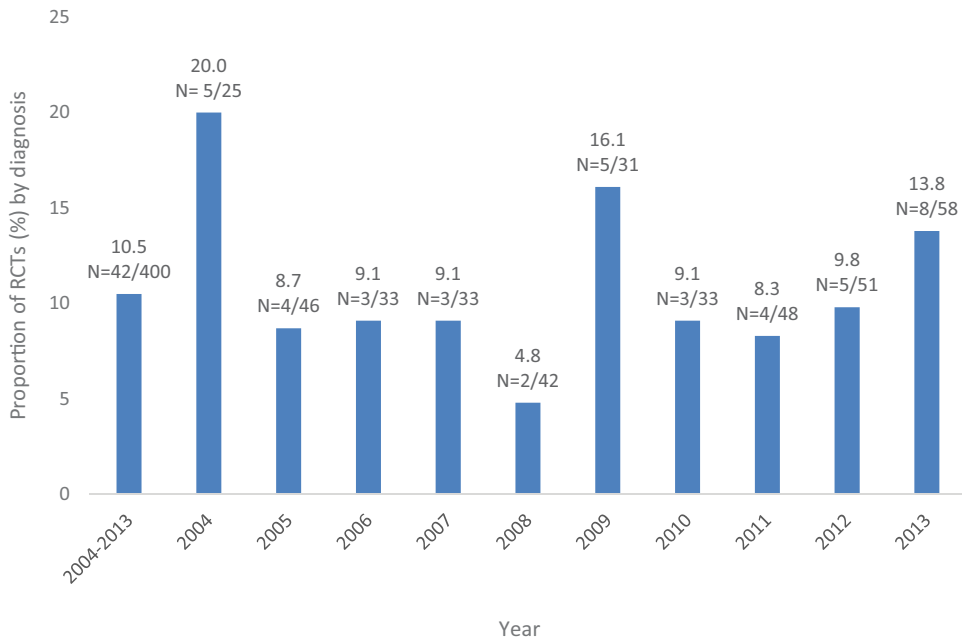


**Fig. 2** Proportion of randomized controlled trials (RCTs) as a surrogate for quality of evidence regarding lumbar fusion increasing over the past 10 years.

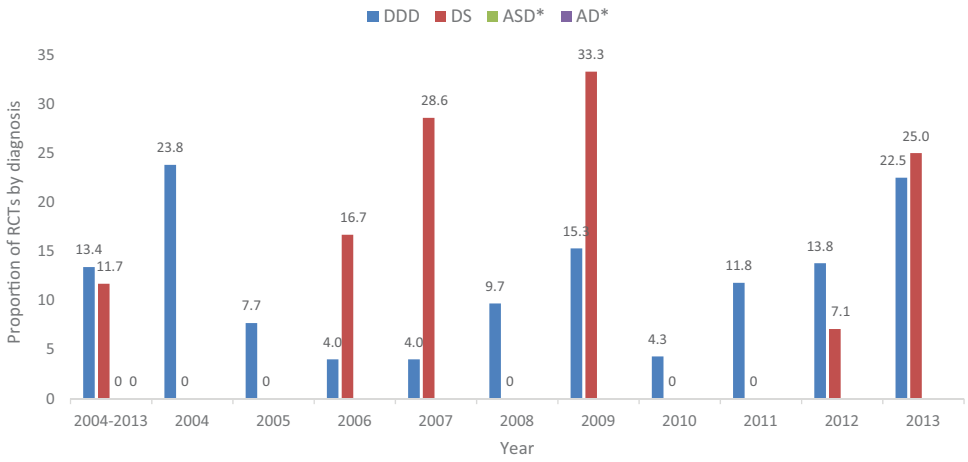


**Fig. 3** The difference in the proportion of RCTs among the four primary fusion diagnoses (DDD, DS, ASD, AD) over the past 10 years. Abbreviations: AD, adult deformity; ASD, adjacent segment disease; DDD, degenerative disk disease; DS, degenerative spondylolisthesis; RCT, randomized controlled trail. *No RCT found evaluating ASD or AD.

- The smallest proportion of fusion RCTs evaluating patients with DDD occurred in the years 2006 ($n = 1/25$; 4%), 2007 ($n = 1/25$; 4%), and 2010 ($n = 1/23$; 4.3%).
- The greatest proportion of fusion RCTs evaluating patients with DS occurred in the year 2009 ($n = 1/3$; 33.3%) followed by the years 2007 ($n = 2/7$; 28.6%) and 2013 ($n = 2/8$; 25.0%).
- Five years (2004, 2005, 2008, 2010, and 2011) did not include any RCTs evaluating patients with DS.
- The proportion of fusion RCTs evaluating patients with DS for the remaining 2 years was 2006 (16.7%) and 2007 (7.1%).

### Clinical Question 3: Is There a Difference in Type and Quality of Clinical Outcomes Measured among RCTs over Time? (►Fig. 4)

- Of the 42 included RCTs, 37 trials (88.1%) included patient-reported outcomes, 16 (38.1%) reported on clinician-based outcomes, and two studies (4.8%) did not report type of outcomes.
- Thirty-three studies (78.6%) administered the Oswestry Disability Index, 25 studies (59.5%) administered a pain visual analog scale, and 17 studies (40.5%) administered the Short-Form 36 (►**Fig. 4**).
- There was no trend over time regarding type or quality of outcome.

### Clinical Question 4: Is There a Difference in Type and Quality of Adverse Events Measured among RCTs over Time?

- Of the 42 included RCTs, 34 trials (81%) included complications, 25 (59.5%) included reoperations, and 5 (11.9%) did not report any adverse events.
- The most common adverse events reported across the studies were reoperation (59.5%), dural sac tear (26.2%), and deep vein thrombosis (16.7%).
- There were 5 trials (11.9%) that included an adverse events severity system.
- There was no trend over time regarding adverse events severity system as these 5 trials were from the years 2004, 2008, 2009, 2010, and 2011.

### Clinical Question 5: Are There Changes in Fusion Treatment Approaches over Time by Diagnosis over the Past 10 Years?

- Over the course of the 10-year period, anterior, posterior, circumferential, transforaminal, and a combination of these approaches have been used.
- A posterior approach was used in 33.3%; circumferential in 21.4%; anterior in 19%; transforaminal in 11.9%; combination of one or more approaches used in 9.5%; and one study did not report a specific approach (2.4%).
- There were no discernible changes in treatment approaches over time or by diagnosis in the past 10 years.

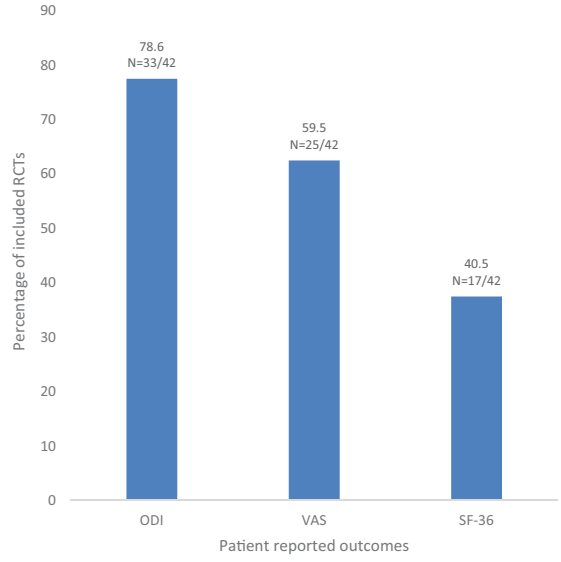


**Fig. 4** Percentage of included randomized controlled trials (RCTs) measuring Oswestry Disability Index (ODI), visual analog scale (VAS), and Short-Form 36 (SF-36).

**Table 1** Demographics and characteristic of included studies

| Investigator | Year | Industry funded | *n* | % male | Age, y (mean ± SD) | Diagnosis | Fusion approach | Comparison | Outcome type | Outcome measures | Adverse events type | Adverse events severity scoring | Adverse events |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Guyer[25] | 2004 | Yes | 144 | NR | NR | DDD | AF | TDR | PRO | VAS; ODI | Complications; reoperation | No | Heterotopic ossification; retrograde ejaculation; bowel obstruction; depression; adynamic ileus; infection; degenerative changes; neurologic deficits; reoperation |
| Sasso[41] | 2004 | Yes | 140 | 45.3 | 41 | DDD | AF+ | AF+ | PRO; CBO | ODI; radiographic fusion; low back pain questionnaire; SF-36; neurologic status; overall health | Complications | No | Vascular intraoperative; pain; neurologic; incisional; spinal event; urologic; gastrointestinal; retrograde ejaculation; respiratory; trauma; peritoneal; vascular postoperative; bone fracture; implant displacement; nonunion; meningitis; implant breakage; death |
| Zigler[50] | 2004 | NR | 39 | 51.2 | 38.8 | DDD | CF | TDR | PRO; CBO | VAS for pain; ODI; patient satisfaction; range of motion; activity level | Complications; reoperation | No | Pain; intraoperative complications; reoperation; dislodgement of spacer; iliac vein laceration; infection; deep vein thrombosis |
| Keller[28] | 2004 | No | 124 | 45 | 43 | DDD | PF | CBT | PRO | ODI; muscle strength; Biering-Sorensen test | NR | NR | NR |
| Geisler[21] | 2004 | NR | 304 | 51.6 | 39.6 | DDD | AF | TDR | PRO | ODI; VAS for pain | Complications | Yes | Neurologic |
| Blumenthal[6] | 2005 | Yes | 304 | 51.6 | 39.6 | DDD | AF | TDR | PRO | VAS for pain; ODI; SF-36; neurologic status; patient satisfaction | Complications; reoperation | No | Death; venous injury; sexual dysfunction; ileus; deep vein thrombosis; significant blood loss; hernia; dural tear; arterial thrombosis; infection; pseudarthrosis; donor site pain; subsidence; reoperation |
| Fairbank[18] | 2005 | Yes | 349 | 49.3 | NR | DDD | NR | Exercise | PRO | ODI; shuttle walking test; SF-36; Zung Depression Scale; somatic perception questionnaire | Complications; reoperation | No | Dural tear; excessive bleeding; implant problems; bone fracture; vascular injury; broken drain; hemorrhage; reoperation |

(*Continued*)

**Table 1** (*Continued*)

| Investigator | Year | Industry funded | *n* | % male | Age, y (mean ± SD) | Diagnosis | Fusion approach | Comparison | Outcome type | Outcome measures | Adverse events type | Adverse events severity scoring | Adverse events |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| McAfee[33] | 2005 | Yes | 304 | 51.6 | 39.6 | DDD | AF | TDR | CBO | Range of motion; disk space height | Complications | No | Subsidence |
| McKenna[35] | 2005 | NR | 83 | 44.9 | 40.3 | DDD | CF+ | CF+ | PRO | ODI; VAS for pain; SF-36 | Complications | No | Infection; transient radiculopathy; retrograde ejaculation; donor site pain; vascular injury; dural tear; bowel perforation; wound hematoma; incisional hernia |
| Brox[9] | 2006 | Yes | 60 | 52 | 42.5 | DDD | PF | CBT | PRO; CBO | ODI; VAS for pain; general function score; Hopkins Emotional Distress Score; fear-avoidance belief questionnaire; life satisfaction; global back disability question; Prolo scale; work status; finger-tip-floor distance | Complications | No | Infection |
| McAfee[34] | 2006 | Yes | 304 | 51.6 | 39.6 | DDD | AF | TDR | NR | NR | Reoperation | No | Reoperation |
| Videbaek[42] | 2006 | No | 148 | 60.2 | 45.5 | DDD | PF | CF | PRO | Dallas pain questionnaire; ODI; SF-36; low back pain rating scale | NR | NR | NR |
| Fernández-Fairen[19] | 2007 | No | 82 | 37.8 | 61.1 | DS | PF+ | PF+ | PRO; CBO | SF-36; radiographic disk height | Complications; reoperation | No | Nerve root irritation; nonunion; reoperation |
| Zigler[49] | 2007 | No | 236 | 56.5 | 41.8 | DDD | CF | TDR | PRO; CBO | ODI; SF-36; VAS for pain; VAS for satisfaction; neurologic success; radiologic outcomes; narcotic use; work status; recreation status | Complications | No | Significant blood loss; retrograde ejaculation; infection; deep vein thrombosis |
| Weinstein[43] | 2007 | Yes | 304 | 34 | 66.0 ± 10.0 | DS | PF | Exercise | PRO | ODI; SF-36; Stenosis Bothersome Index; Low Back Pain Bothersome Index; self-reported improvement; self-reported satisfaction | Complications; reoperation | No | Blood loss; dural tear; cerebrospinal fluid leak; vascular injury; nerve root injury; wound infection; death; recurrent stenosis; reoperation |
| Geisler[22] | 2008 | Yes | 375 | 44.2 | 39.3 | DDD | AF | TDR | PRO | ODI; VAS for pain; patient satisfaction | Complications; reoperation | No | Infection; subsidence; implant displacement; neurologic; DDD progression; pain; vessel damage; reoperation |

**Table 1** (*Continued*)

| Investigator | Year | Industry funded | *n* | % male | Age, y (mean ± SD) | Diagnosis | Fusion approach | Comparison | Outcome type | Outcome measures | Adverse events type | Adverse events severity scoring | Adverse events |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sasso[40] | 2008 | Yes | 67 | 49.2 | 38 | DDD | CF | TDR | PRO; CBO | ODI; VAS; radiographic motion | Complications; reoperation | Yes | Infection; pain; hematoma, end plate fracture; hardware migration; vascular injury; reoperation; tachyarrhythmia; hypoxia; pulmonary embolism; extraperitoneal seroma |
| Berg[4] | 2009 | NR | 152 | 40.8 | 39.4 ± 8.0 | DDD | PF | TDR | PRO | Global assessment; VAS for pain; ODI; SF-36; Eq. 5D; patient satisfaction; work status | Complications; reoperation | No | Sexual dysfunction; reoperation |
| Guyer[24] | 2009 | Yes | 133 | 53.4 | 39.6 | DDD | AF | TDR | PRO; CBO | VAS for pain; ODI; SF-36; patient satisfaction; radiographic range of motion; disk height; segmental translation; work status | Complications; reoperation | No | Depression; adynamic ileus |
| Auerbach[3] | 2009 | Yes | 200 | 52.5 | 39 | DDD | CF | TDR | CBO | Radiographic outcomes | NR | NR | NR |
| Berg[5] | 2009 | NR | 152 | 40.8 | 39.4 ± 8.0 | DDD | PF | TDR | PRO | Global assessment; VAS for pain; ODI; SF-36; Eq. 5D; patient satisfaction; work status | Complications; reoperation | Yes | Infection; hematoma; facet joint problem; pseudarthrosis; hernia; nerve entrapment; donor site pain; adjacent segment disease; dural tear; meralgia paresthetica; subsidence; reoperation |
| Weinstein[44] | 2009 | Yes | 304 | 34 | 66.0 ± 10.0 | DS | PF | Exercise | PRO | ODI; SF-36; Stenosis Bothersome Index; Low Back Pain Bothersome Index; self-reported improvement; self-reported satisfaction | Complications; reoperation | No | Blood loss; dural tear; cerebrospinal fluid leak; vascular injury; nerve root injury; wound infection; death; recurrent stenosis; reoperation |
| Brox[8] | 2010 | NR | 124 | NR | NR | DDD | PF | CBT | PRO; CBO | ODI; VAS for pain; general function score; Hopkins Emotional Distress Score; fear-avoidance belief questionnaire; life satisfaction; global back disability question; Prolo scale; | Complications; reoperation | No | Infection; death; reoperation |

(*Continued*)

**Table 1** (*Continued*)

| Investigator | Year | Industry funded | *n* | % male | Age, y (mean ± SD) | Diagnosis | Fusion approach | Comparison | Outcome type | Outcome measures | Adverse events type | Adverse events severity scoring | Adverse events |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | work status; finger-tip-floor distance | | | |
| Putzier[38] | 2010 | NR | 60 | 51.7 | 44 | DDD | CF | Dynamic fixation | PRO; CBO | ODI; VAS for pain; satisfaction; radiologic assessment; pain; functional outcome | Complications | No | Progression of ASD; fusion of dynamically fixated segment; implant failure |
| Ohnmeiss[36] | 2010 | No | 155 | 54.8 | 41.4 (range 19–60) | DDD | AF; PF | TDR | PRO | ODI; VAS for pain; overall satisfaction | Complications | Yes | Nausea; constipation; falls; pain; cancer |
| Delamarter[13] | 2011 | Yes | 237 | 56.5 | 41.8 | DDD | CF | TDR | PRO | ODI; SF-36; VAS for pain; VAS for satisfaction; neurologic success; radiographic outcomes; narcotic use; work status; recreation status | Complications; reoperation | No | Dural tear; significant blood loss; deep vein thrombosis |
| Froholdt[20] | 2011 | Yes | 55 | 41.8 | 42.8 | DDD | PF | CBT | PRO | General function score | Reoperation | No | Reoperation |
| Ohtori[37] | 2011 | No | 41 | 58.5 | 34 | DDD | AF; PF | Exercise | PRO | ODI; JOA; VAS for pain; self-reported subjective outcome | NR | NR | NR |
| Gornet[23] | 2011 | Yes | 577 | 50.4 | NR (range 18–70) | DDD | AF | TDR | PRO | ODI; SF-36; numeric rating scale for pain; patient satisfaction; global perceived effect; work status | Complications; reoperation | Yes | Pain; infection; depression; death; implant displacement; cardiovascular; neurologic; nonunion; vascular injury; peritoneal tear; vertebral fracture; subsidence; allergic reaction; reoperation |
| Aoki[2] | 2012 | NR | 50 | 40 | 65.9 ± 8.8 | DS | TF+ | TF+ | PRO | VAS; JOA | Complications; reoperation | No | Cage migration; nerve root irritation; pulmonary embolism; dural tear; reoperation |
| Xie[45] | 2012 | Yes | 108 | 44.4 | 55.6 | DDD | PF+ | PF+ | PRO | JOA; SF-36 | Complications | No | Infection; dural tear; motor weakness |
| Xue[46] | 2012 | No | 80 | 43.8 | 57.7 | DDD | TF+ | TF+ | PRO | VAS; ODI; Prolo | Complications; reoperation | No | Infection; cerebrospinal fluid leak; deep vein thrombosis; screw failure; reoperation |
| Zigler[51] | 2012 | No | 236 | 56.5 | 41.8 | DDD | CF | TDR | PRO; CBO | VAS for pain; ODI; VAS for satisfaction; range of motion; activity level | Complications; reoperation | No | Excessive blood loss; dural tear; retrograde ejaculation; infection; deep vein thrombosis; death; reoperation |

**Table 1** (*Continued*)

| Investigator | Year | Industry funded | *n* | % male | Age, y (mean ± SD) | Diagnosis | Fusion approach | Comparison | Outcome type | Outcome measures | Adverse events type | Adverse events severity scoring | Adverse events |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Zigler[52] | 2012 | No | 236 | 56.5 | 41.8 | DDD | CF | TDR | CBO | Radiographic changes | NR | NR | NR |
| Choi[11] | 2013 | NR | 54 | 39.6 | 54.8 | DDD | TF+ | TF+ | PRO | ODI; VAS | Complications; reoperation | No | Cage migration; disk herniation; reoperation |
| Davis[12] | 2013 | Yes | 322 | NR | NR | DS | PF | Interlaminar stabilization | PRO; CBO | ODI; VAS; SF-12; Zurich Claudication questionnaire; FDA overall fusion success; quantitative radiographic data | Complications; reoperation | No | Spinous process fracture |
| Duncan[17] | 2013 | NR | 102 | 39.2 | 54.7 | DDD | TF+ | TF+ | NR | NR | Complications | No | Cage migration |
| Høy[26] | 2013 | NR | 100 | 41 | 50 | DDD; DS | PF | TF | PRO | Dallas Pain Questionnaire; ODI; SF-36; low back pain rating scale; daily activity; work leisure; anxiety/depression; social interest | Complications; reoperation | No | Hematoma; infection; nerve root lesion; dural tear; pneumothorax; implant failure; reoperation |
| Zhang[48] | 2013 | Yes | 68 | 35.2 | 57.5 | DDD; DS | TF+ | TF+ | PRO; CBO | VAS; ODI; SF-36; radiographic | Complications | No | Tube; urinary tract infection; epididymitis; lateral epicondylitis |
| Lin[29] | 2013 | No | 85 | 45.8 | 66.3 | DDD | AF+ | AF+ | PRO | ODI; VAS | Complications | No | Foot drop |
| Liu[30] | 2013 | No | 120 | 25.8 | 58.3 | DDD | PF+ | PF+ | PRO; CBO | JOA; radiographic measures | Complications; reoperation | No | Six cases of transient neurologic deficits |
| Mannion[32] | 2013 | Yes | 473 | | 41.2 (± 8.3) | DDD | PF | CBT | PRO | ODI; VAS for pain; medication use; work status; EuroQol; VAS for HRQOL; VAS for satisfaction; VAS for global treatment | Reoperation | No | Reoperation |

Abbreviations: AF, anterior fusion; CBO, clinician-based outcome; CF, circumferential fusion; CBT, cognitive behavior therapy; DDD, degenerative disk disease; DS, degenerative spondylolisthesis; EuroQoL, European quality of life; JOA, Japanese Orthopedic Association; NR, not reported; ODI, Oswestry Disability Index; PF, posterior fusion; PRO, patient-reported outcome; TF, transforaminal fusion; TDR, total disk replacement; SF-36, Short-Form 36; VAS, visual analog scale; +, fusion + hardware comparison.

## Discussion

This structured review was performed in an effort to assess whether the quality of clinical research on lumbar fusion has shown consistent improvement over the past decade. In the end, we are unable to make clear statements regarding trends over this period. On the other hand, there are some positive features to be noted from our results.

Although there has not been an apparent shift toward a greater percentage of RCT design among published studies, there has been a steady increase in the number of RCT studies published with a focus on DDD and on DS. As the two most common surgical indications for fusion, it is an encouraging finding. Although it is beyond the scope of this article to derive treatment guidelines, the numbers available suggest that there has likely emerged a relatively high level of evidence data on which to base such recommendations.

We are also encouraged by the relatively high percentage (88.1%) of RCTs using validated, patient-centered outcomes over the past decade. The most widely used questionnaire was the Oswestry Disability Index, which was used in 78.6% of reviewed RCTs. Although debate regarding which outcomes instruments are the best designed or the most responsive for patients receiving lumbar fusion is perhaps unsettled, the importance of using validated, patient-reported outcomes as opposed to clinician-reported outcomes is well accepted. This approach appears to be fairly consistently used by authors of the highest level of medical evidence in the field of lumbar fusion.

Unfortunately, the same cannot be said regarding the reportage of adverse events in these same studies. Although 81% of RCTs did include some discussion of adverse events, only 11.9% utilized some classification or scale of complications, which may in part reflect the lack of availability or development of clinical research tools with a valid weighting of adverse events following lumbar fusion surgery. We hope that this review may serve as an illustration of the need for such an effort.

The lack of a consistent approach to surgical fusion remains a barrier to development of a reliable body of high-quality clinical data on which to base treatment recommendations. Although the variety of approaches available does reflect a significant effort and investment in surgical innovation, it is unlikely that all of the approaches currently in use are equally safe or effective. Although undoubtedly some clinical decision making regarding approach is tailored to the needs of an individual patient, it is also likely driven at least in part by the training and experience of the surgeon performing the procedure.[27] This review highlights the need for higher-level comparisons of specific surgical approaches and techniques.

The lack of high-level data to assess fusion for patients with adult spinal deformity or adjacent segment disease remains an area of concern. The lack of published RCTs in these areas may reflect the even greater variations of clinical presentation and surgical approach among such patients. The comparatively smaller number of such patients also presents difficulty in obtaining patient cohorts of sufficient size to allow meaningful statistical comparisons. Despite such obstacles, however, patients and surgeons would undoubtedly benefit from efforts at improving the clinical data guiding treatment recommendations.

This review ultimately does not prove that the quality of the reported data is truly improved. A more detailed analysis of the actual content of the published studies would be required to gain a better understanding of their true level of quality. Nonetheless, this study does provide at least a partial assessment of the current landscape of lumbar spine clinical research. Our results do show that there appears to be an increasing adoption of an EBM-supported approach within the discipline of lumbar spine surgery over the past decade.

## Conclusion

This structured review demonstrates that there has been an increase in the available clinical database of RCTs using patient-reported outcomes evaluating the benefit of lumbar spinal fusion for the diagnoses of DDD and DS. Gaps remain in the standardization of reportage of adverse events in such trials, as well as uniformity of surgical approaches used. Finally, continued efforts to develop higher-quality data for other surgical indications for lumbar fusion, most notably in the presence of adult spinal deformity and revision of prior surgical fusions, appear warranted.

Disclosures
Robert Hart, Board membership: CSRS, ISSLS, ISSGF; Consultant: DePuy Spine, Globus, Medtronic; Royalties: Seaspine, DePuy Synthes
Jeffrey T. Hermsmeyer, none
Rajiv K. Sethi, none
Daniel C. Norvell, none

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

## References

1 Anderson PA, Hart RA. Adverse events recording and reporting in clinical trials of cervical total disk replacement. Instr Course Lect 2014;63:287–296
2 Aoki Y, Yamagata M, Ikeda Y, et al. A prospective randomized controlled study comparing transforaminal lumbar interbody fusion techniques for degenerative spondylolisthesis: unilateral pedicle screw and 1 cage versus bilateral pedicle screws and 2 cages. J Neurosurg Spine 2012;17(2):153–159
3 Auerbach JD, Jones KJ, Milby AH, Anakwenze OA, Balderston RA. Segmental contribution toward total lumbar range of motion in disc replacement and fusions: a comparison of operative and adjacent levels. Spine (Phila Pa 1976) 2009;34(23):2510–2517
4 Berg S, Fritzell P, Tropp H. Sex life and sexual function in men and women before and after total disc replacement compared with posterior lumbar fusion. Spine J 2009;9(12):987–994
5 Berg S, Tullberg T, Branth B, Olerud C, Tropp H. Total disc replacement compared to lumbar fusion: a randomised controlled trial with 2-year follow-up. Eur Spine J 2009;18(10):1512–1519
6 Blumenthal S, McAfee PC, Guyer RD, et al. A prospective, randomized, multicenter Food and Drug Administration investigational device exemptions study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part I: evaluation of clinical outcomes. Spine (Phila Pa 1976) 2005;30(14): 1565–1575, discussion E387–E391

7 Bonaldi G, Brembilla C, Cianfoni A. Minimally-invasive posterior lumbar stabilization for degenerative low back pain and sciatica. A review. Eur J Radiol 2015;84(5):789–798
8 Brox JI, Nygaard OP, Holm I, Keller A, Ingebrigtsen T, Reikerås O. Four-year follow-up of surgical versus non-surgical therapy for chronic low back pain. Ann Rheum Dis 2010;69(9):1643–1648
9 Brox JI, Reikerås O, Nygaard Ø, et al. Lumbar instrumented fusion compared with cognitive intervention and exercises in patients with chronic back pain after previous surgery for disc herniation: a prospective randomized controlled study. Pain 2006;122(1-2): 145–155
10 Carragee EJ, Hurwitz EL, Weiner BK. A critical review of recombinant human bone morphogenetic protein-2 trials in spinal surgery: emerging safety concerns and lessons learned. Spine J 2011; 11(6):471–491
11 Choi UY, Park JY, Kim KH, et al. Unilateral versus bilateral percutaneous pedicle screw fixation in minimally invasive transforaminal lumbar interbody fusion. Neurosurg Focus 2013;35(2):E11
12 Davis RJ, Errico TJ, Bae H, Auerbach JD. Decompression and Coflex interlaminar stabilization compared with decompression and instrumented spinal fusion for spinal stenosis and low-grade degenerative spondylolisthesis: two-year results from the prospective, randomized, multicenter, Food and Drug Administration Investigational Device Exemption trial. Spine (Phila Pa 1976) 2013;38(18):1529–1539
13 Delamarter R, Zigler JE, Balderston RA, Cammisa FP, Goldstein JA, Spivak JM. Prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement compared with circumferential arthrodesis for the treatment of two-level lumbar degenerative disc disease: results at twenty-four months. J Bone Joint Surg Am 2011;93(8):705–715
14 Deyo RA. Fusion surgery for lumbar degenerative disc disease: still more questions than answers. Spine J 2015;15(2):272–274
15 Deyo RA, Dworkin SF, Amtmann D, et al. Focus article report of the NIH task force on research standards for chronic low back pain. Clin J Pain 2014;30(8):701–712
16 Deyo RA, Gray DT, Kreuter W, Mirza S, Martin BI. United States trends in lumbar fusion surgery for degenerative conditions. Spine (Phila Pa 1976) 2005;30(12):1441–1445, discussion 1446–1447
17 Duncan JW, Bailey RA. An analysis of fusion cage migration in unilateral and bilateral fixation with transforaminal lumbar interbody fusion. Eur Spine J 2013;22(2):439–445
18 Fairbank J, Frost H, Wilson-MacDonald J, Yu LM, Barker K, Collins R; Spine Stabilisation Trial Group. Randomised controlled trial to compare surgical stabilisation of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: the MRC spine stabilisation trial. BMJ 2005;330(7502):1233
19 Fernández-Fairen M, Sala P, Ramírez H, Gil J. A prospective randomized study of unilateral versus bilateral instrumented posterolateral lumbar fusion in degenerative spondylolisthesis. Spine (Phila Pa 1976) 2007;32(4):395–401
20 Froholdt A, Holm I, Keller A, Gunderson RB, Reikeraas O, Brox JI. No difference in long-term trunk muscle strength, cross-sectional area, and density in patients with chronic low back pain 7 to 11 years after lumbar fusion versus cognitive intervention and exercises. Spine J 2011;11(8):718–725
21 Geisler FH, Blumenthal SL, Guyer RD, et al. Neurological complications of lumbar artificial disc replacement and comparison of clinical results with those related to lumbar arthrodesis in the literature: results of a multicenter, prospective, randomized investigational device exemption study of Charité intervertebral disc. Invited submission from the Joint Section Meeting on Disorders of the Spine and Peripheral Nerves, March 2004. J Neurosurg Spine 2004;1(2):143–154
22 Geisler FH, Guyer RD, Blumenthal SL, et al. Effect of previous surgery on clinical outcome following 1-level lumbar arthroplasty. J Neurosurg Spine 2008;8(2):108–114
23 Gornet MF, Burkus JK, Dryer RF, Peloza JH. Lumbar disc arthroplasty with Maverick disc versus stand-alone interbody fusion: a prospective, randomized, controlled, multicenter investigational device exemption trial. Spine (Phila Pa 1976) 2011;36(25):E1600–E1611
24 Guyer RD, McAfee PC, Banco RJ, et al. Prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: five-year follow-up. Spine J 2009;9(5):374–386
25 Guyer RD, McAfee PC, Hochschuler SH, et al. Prospective randomized study of the Charite artificial disc: data from two investigational centers. Spine J 2004;4(6, Suppl):252S–259S
26 Høy K, Bünger C, Niederman B, et al. Transforaminal lumbar interbody fusion (TLIF) versus posterolateral instrumented fusion (PLF) in degenerative lumbar disorders: a randomized clinical trial with 2-year follow-up. Eur Spine J 2013;22(9):2022–2029
27 Irwin ZN, Hilibrand A, Gustavel M, et al. Variation in surgical decision making for degenerative spinal disorders. Part I: lumbar spine. Spine (Phila Pa 1976) 2005;30(19):2208–2213
28 Keller A, Brox JI, Gunderson R, Holm I, Friis A, Reikerås O. Trunk muscle strength, cross-sectional area, and density in patients with chronic low back pain randomized to lumbar fusion or cognitive intervention and exercises. Spine (Phila Pa 1976) 2004;29(1):3–8
29 Lin B, Xu Y, He Y, Zhang B, Lin Q, He M. Minimally invasive unilateral pedicle screw fixation and lumbar interbody fusion for the treatment of lumbar degenerative disease. Orthopedics 2013;36(8):e1071–e1076
30 Liu H, Wu W, Li Y, Liu J, Yang K, Chen Y. Protective effects of preserving the posterior complex on the development of adjacent-segment degeneration after lumbar fusion: clinical article. J Neurosurg Spine 2013;19(2):201–206
31 Lubelski D, Mihalovich KE, Skelly AC, et al. Is minimal access spine surgery more cost-effective than conventional spine surgery? Spine (Phila Pa 1976) 2014;39(22, Suppl 1):S65–74
32 Mannion AF, Brox JI, Fairbank JC. Comparison of spinal fusion and nonoperative treatment in patients with chronic low back pain: long-term follow-up of three randomized controlled trials. Spine J 2013;13(11):1438–1448
33 McAfee PC, Cunningham B, Holsapple G, et al. A prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part II: evaluation of radiographic outcomes and correlation of surgical technique accuracy with clinical outcomes. Spine (Phila Pa 1976) 2005;30(14):1576–1583, discussion E388–E390
34 McAfee PC, Geisler FH, Saiedy SS, et al. Revisability of the CHARITE artificial disc replacement: analysis of 688 patients enrolled in the U.S. IDE study of the CHARITE Artificial Disc. Spine (Phila Pa 1976) 2006;31(11):1217–1226
35 McKenna PJ, Freeman BJ, Mulholland RC, Grevitt MP, Webb JK, Mehdian SH. A prospective, randomised controlled trial of femoral ring allograft versus a titanium cage in circumferential lumbar spinal fusion with minimum 2-year clinical results. Eur Spine J 2005;14(8):727–737
36 Ohnmeiss DD, Bodemer W, Zigler JE. Effect of adverse events on low back surgery outcome: twenty-four-month follow-up results from a Food and Drug Administration investigational device exemptiontrial. Spine (Phila Pa 1976) 2010;35(7):835–838
37 Ohtori S, Koshi T, Yamashita M, et al. Surgical versus nonsurgical treatment of selected patients with discogenic low back pain: a small-sized randomized trial. Spine (Phila Pa 1976) 2011;36(5):347–354
38 Putzier M, Hoff E, Tohtz S, Gross C, Perka C, Strube P. Dynamic stabilization adjacent to single-level fusion: part II. No clinical benefit for asymptomatic, initially degenerated adjacent segments after 6 years follow-up. Eur Spine J 2010;19(12):2181–2189
39 Robinson JC. Hospital market concentration, pricing, and profitability in orthopedic surgery and interventional cardiology. Am J Manag Care 2011;17(6 Spec No.):e241–e248

40 Sasso RC, Foulk DM, Hahn M. Prospective, randomized trial of metal-on-metal artificial lumbar disc replacement: initial results for treatment of discogenic pain. Spine (Phila Pa 1976) 2008;33(2):123–131
41 Sasso RC, Kitchel SH, Dawson EG. A prospective, randomized controlled clinical trial of anterior lumbar interbody fusion using a titanium cylindrical threaded fusion device. Spine (Phila Pa 1976) 2004;29(2):113–122, discussion 121–122
42 Videbaek TS, Christensen FB, Soegaard R, et al. Circumferential fusion improves outcome in comparison with instrumented posterolateral fusion: long-term results of a randomized clinical trial. Spine (Phila Pa 1976) 2006;31(25):2875–2880
43 Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. N Engl J Med 2007;356(22):2257–2270
44 Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical compared with nonoperative treatment for lumbar degenerative spondylolisthesis. four-year results in the Spine Patient Outcomes Research Trial (SPORT) randomized and observational cohorts. J Bone Joint Surg Am 2009;91(6):1295–1304
45 Xie Y, Ma H, Li H, et al. Comparative study of unilateral and bilateral pedicle screw fixation in posterior lumbar interbody fusion. Orthopedics 2012;35(10):e1517–e1523
46 Xue H, Tu Y, Cai M. Comparison of unilateral versus bilateral instrumented transforaminal lumbar interbody fusion in degenerative lumbar diseases. Spine J 2012;12(3):209–215
47 Yoshihara H, Yoneoka D. National trends in the surgical treatment for lumbar degenerative disc disease: United States, 2000 to 2009. Spine J 2015;15(2):265–271
48 Zhang K, Sun W, Zhao CQ, et al. Unilateral versus bilateral instrumented transforaminal lumbar interbody fusion in two-level degenerative lumbar disorders: a prospective randomised study. Int Orthop 2014;38(1):111–116
49 Zigler J, Delamarter R, Spivak JM, et al. Results of the prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement versus circumferential fusion for the treatment of 1-level degenerative disc disease. Spine (Phila Pa 1976) 2007;32(11):1155–1162, discussion 1163
50 Zigler JE. Lumbar spine arthroplasty using the ProDisc II. Spine J 2004;4(6, Suppl):260S–267S
51 Zigler JE, Delamarter RB. Five-year results of the prospective, randomized, multicenter, Food and Drug Administration investigational device exemption study of the ProDisc-L total disc replacement versus circumferential arthrodesis for the treatment of single-level degenerative disc disease. J Neurosurg Spine 2012;17(6):493–501
52 Zigler JE, Glenn J, Delamarter RB. Five-year adjacent-level degenerative changes in patients with single-level disease treated using lumbar total disc replacement with ProDisc-L versus circumferential fusion. J Neurosurg Spine 2012;17(6):504–511

# Editorial Perspective

*EBSJ* reviewers welcomed this systematic review by Hart and coauthors. There were several important commentaries regarding this study that *EBSJ* wanted to share with our readership:

- The premise that prospective randomized clinical trials (PRCTs) represent the height of scientific evidence in surgical care has become something that has been increasingly challenged (see Editorial "Nothing Hurts Follow-Up like Follow-Up" on page 165 of this issue). A PRCT studies "efficacy" of a procedure—it seeks to prove or disprove the likelihood of a given intervention in comparison to another treatment to result in a desired therapeutic effect under tightly controlled circumstances. The purported main benefit of this type of "explanatory" RCT is the promise of bias reduction. In light of an apparent increasing unwillingness of some populations to allow their care to be chosen by randomization—even under the premise of therapeutic equipoise—the role of efficiency trials, meaning studies where treatments are studied in a real life practice of medicine, has gained increasing consideration. It is not difficult to foresee where large-scale "pragmatic trials" and registry-derived studies may supersede surgical PRCTs as the most impactful study on the evidence pyramid. Therefore, the current study premise of the authors to focus on level 1 PRCTs as the pinnacle of scientific validity may not be representative of the actually most meaningful form of research for the future.
- The current study has further underscored the ongoing categorical confusion of studies using the clinical symptom of "low back pain" as their study foundation. Indeed many studies lump together entities such as such as "discogenic back pain," "degenerative spondylolisthesis," "(postdiskectomy) disk degeneration," and "stable'" (isthmic) spondylolisthesis based on their common generalized clinical presentation of "low back pain." Part of this confusion arises out of our lack of universally accepted operational definitions. Part of the problem also arises out of the insufficient specificity of the International Classification of Diseases, 9th Revision system with its overabundance of spine related terms. The reviewers expressed the hope that the increasing prevalence of International Classification of Diseases, 10th Revision and electronic medical records will foster improved specificity of medical terminology. The use of undifferentiated terms such as "low back pain" as presenting symptomatology without subdifferentiation for inclusion in PRCTs will likely not be sustainable in the future.
- One reviewer pointed out the ongoing common disregard of nonorganic factors in studies regarding back pain. Clinical comorbidities such as anxiety, depression, fear avoidance, catastrophizing, presence of pre-existing chronic pain, sleep deprivation, and many other psychosocial variables likely heavily influence patient-reported outcomes more than the actual treatment interventions, thus leading to spurious result reporting.
- In conclusion, the reviewers welcomed the finding of an increasing number of PRCTs being generated on the subject of lumbar fusions but warned of placing too much emphasis on PRCTs in generalized discussions regarding preferred treatments of "low back pain" without necessary further differentiation and due deliberation of "treatment efficiency." Finally, the reviewers shared the authors' surprise that there had been no high-level studies on the subject of adult degenerative scoliosis and adjacent segment disease.