Spine Deformity
www.spine-deformity.org

Spine Deformity 3 (2015) 312–317

Systematic Review

Patient Reported Outcomes in Adult Spinal Deformity Surgery: A Bibliometric Analysis

Holt S. Cutler, BSE[a], Javier Z. Guzman, BS[a], Motasem Al Maaieh, MD[a], James Connolly, BA[a], Branko Skovrlj, MD[b], Samuel K. Cho, MD[a,*]

[a]*Department of Orthopaedic Surgery, Icahn School of Medicine at Mount Sinai, 1428 Madison Avenue, New York, NY 10029, USA*
[b]*Department of Neurosurgery, Icahn School of Medicine at Mount Sinai, 1428 Madison Avenue, New York, NY 10029, USA*

Received 23 June 2014; revised 11 October 2014; accepted 14 December 2014

Abstract

Study design: Bibliometric analysis.
Objectives: To identify patient-reported outcomes (PROs) used in adult spinal deformity (ASD) research over the past decade, their frequency, and usage trends.
Summary of background data: The emphasis on PROs is increasing along with the demand for evidence-based medicine. However, there is currently no standardization or consensus on which PROs ought to be used in ASD.
Methods: Five top orthopedics journals were reviewed from 2004 to 2013 for clinical studies of surgical intervention in ASD that report PROs. Publication year, level of evidence (LOE), and PROs were collected for each article. Errors and inconsistencies of PRO score reporting were analyzed for the 3 most commonly used PROs.
Results: A total of 84 PRO studies were published in ASD literature over the period studied. The number of PRO studies published increased from 1 in 2004 to 16 in 2013. We identified 24 unique PROs. The Oswestry Disability Index (ODI) was the most frequently used single instrument (47.8%), followed by the Scoliosis Research Society (SRS)-22 (35.6%) and SRS-24 (21.1%), and Short Form-36 (SF-36) and visual analog scale (VAS) were tied (13.3%). The combined use of SRS instruments exceeded ODI use. LOE 4 was most common (42.9%), and no LOE 1 studies were identified. Incomplete preoperative and postoperative PRO scores was the most common reporting inconsistency, occurring in 16% of articles using ODI, 58% of articles using SRS-24, and 22% of articles using SRS-22.
Conclusions: The frequency of studies using PROs in ASD research has increased over the past decade, yet quality studies and standardization are lacking. In general, the ODI and SRS instruments are emerging as standards in ASD surgery; however, frequent use of many uncommon PROs presents a challenge for interstudy comparisons. Additionally, of the top 5 instruments used, only SF-36 is routinely used for cost-effectiveness studies, making procedure cost–outcome decisions difficult.

Keywords: Adult spinal deformity; Patient-reported outcomes; Oswestry; Scoliosis research society; Health-related quality of life

Introduction

Studies of outcomes in adult spinal deformity surgery (ASD) demonstrate treatment effectiveness and enable healthcare providers and patients to make informed decisions. Over the past decade, the cost and utilization of spinal deformity surgery, and other surgical fields has risen dramatically prompting hospitals and regulatory bodies to push for cost-effective and evidence-based treatments. In surgically managed ASD patients, McCarthy et al. [1] estimated a 16-fold increase in total charges in the Medicare population from 2000 to 2010. The increase is due, in part, to the aging baby boomer population driving an increase in utilization. Increases are also attributed to the introduction of new surgical procedures. Outcome studies will provide the necessary data to drive utilization of the most effective spinal deformity treatments and ensure that treatment quality is maintained in the face of rising economic pressures for greater efficiency [1].

Disclosure: HSC (None); JZG (None); MAM (None); JC (None) BS (None) SKC (None).

*Corresponding author. Samuel K. Cho, MD, Department of Orthopaedic Surgery, Icahn School of Medicine at Mount Sinai, 1428 Madison Avenue, New York, NY 10029, USA. Tel.: (212) 241-0276; fax: (212) 534-5841.

E-mail address: samuel.cho@mountsinai.org (S.K. Cho).

2212-134X/$ - see front matter © 2015 Scoliosis Research Society.
http://dx.doi.org/10.1016/j.jspd.2014.12.004

Outcome studies in ASD surgery frequently report complication rates, revision surgery rates, procedure measures, such as curve correction and sagittal or coronal balance, and patient-reported outcomes (PROs). A majority of PROs used and discussed in the literature today are health-related quality of life (HRQOL) instruments, such as the Scoliosis Research Society (SRS)-24 question survey. Simple pain scales, such as the visual analog scale (VAS), predate HRQOL instruments by several decades, but they are still widely used. Robust PRO data are crucial for cost-effectiveness studies as they most closely reflect the fundamental goal of treatment—to improve quality of life for the patient. In an earlier review of SRS and the Oswestry Disability Index (ODI) in the ASD literature, Yadla et al. [2] found very little use of HRQOL instruments prior to 15 years ago. The purpose of our study was to perform a comprehensive review of PRO instrument use in ASD literature, in an effort to highlight trends and usage patterns that may aid researchers deciding between PROs for research studies or professional societies and governing bodies drafting guidelines to standardize PRO use in ASD research.

Methods

Five top orthopedics journals (*The Journal of Bone & Joint Surgery* [*JBJS (Am)*], *The Bone & Joint Journal*[1] [JBJS (Br)], *Spine, European Spine Journal* [ESJ], and *The Spine Journal*) read for information on spine surgery were identified by readership and impact factor. The titles of all clinical articles published in these journals from 2004 through 2013 were reviewed on PubMed. Articles were included if the title referred to a surgical intervention in a clinical study with outcomes measured by any PRO instrument. If these criteria could not be clearly assessed from the title alone, the abstract was reviewed. Articles meeting these criteria were reviewed in full. The title, author, year, level of evidence (LOE), sample size, general diagnosis, and PRO instruments used were recorded. We report the use of several additional PROs beyond those reported in previous literature reviews of ASD, and we distinguish between SRS versions because direct comparisons between them are not valid. After careful review of the abstracts, LOE was assigned to each article by the authors of this study according to the definitions provided by the Oxford Center for Evidence Based Medicine (CEBM) [3]. Only articles having LOE 1 to 4 were included in our study; any LOE 5 articles were excluded.

This study included articles with the general diagnosis of ASD, which included specific diagnoses of adult idiopathic scoliosis, fixed sagittal imbalance, Scheuermann kyphosis, and acquired kyphoscoliotic deformity. Trends in PRO usage were reported over time. The number of PRO studies in each journal was reported along with the frequency of each LOE.

Errors and inconsistencies of PRO score reporting were analyzed by reviewing the Methods and Results sections of all articles that used 1 or more of the 3 most common PROs. The reference standard for reporting was defined as a study with:

- Complete preoperative and postoperative data,
- Numerical data reporting (as opposed to only graphical or statistical reporting),
- Total score reporting (as opposed to only domain, subscale, or component scores), and
- Use of valid, un-modified PRO instruments.

Studies whose PRO reporting did not meet these criteria were identified and categorized according to their errors or inconsistencies.

Data Analysis

Data were recorded and analyzed with Microsoft Excel 2011.

Results

Over the past 10 years (2004–2013), JBJS (Am), JBJS (Br), Spine, ESJ, and The Spine Journal published a total of 19,736 articles. We identified 1,079 clinical studies of surgical interventions that reported use of 1 or more PRO instruments. Of these, 84 articles focused on ASD surgery research. Fig. 1 shows the trend in PRO use over the past decade, with usage increasing from 1 article in 2004 to 16 articles in 2013. Spine published the most PRO studies, with 60 (71.4%) of the 84 total articles in this study, followed by ESJ, The Spine Journal, JBJS (Am), and JBJS (Br). Fig. 2 shows the number of articles published in each journal.

Distinct Outcome Instruments, Frequency, and Level of Evidence

A total of 24 unique PRO instruments were identified in our search. The most common of these are categorized by type in Table 1. The top 5 PROs in order of frequency are

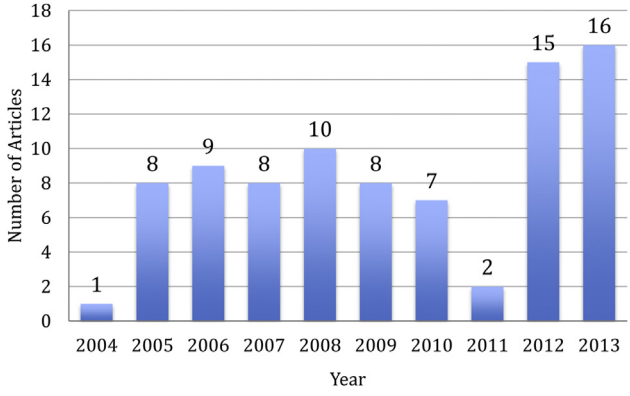

Fig. 1. Patient-reported outcome use by year.

[1] In September 2011, The Journal of Bone and Joint Surgery (British Volume) changed its name to The Bone & Joint Journal. For convenience, articles from both journals are reported under JBJS (Br) in this study.

Table 1
PRO instrument types and top instruments used 2004-2013

General health	Regional specific	Disease specific	Pain
SF-36	ODI	SRS-22r,	VAS
SF-12	RMDQ	SRS-22	NRS
	MacNab Criteria	SRS-24	Wong-Baker FACES
		SRS-29	
		SRS-30	
		JOA Scale	

SF-36, Short Form-36; SF-12, Short Form-12; ODI, Oswestry Disability Index; RMDQ, Roland Morris Disability Questionnaire; SRS, Scoliosis Research Society Questionnaire; JOA, Japanese Orthopaedic Association; VAS, Visual Analog Scale; NRS, Numerical Rating Scale.

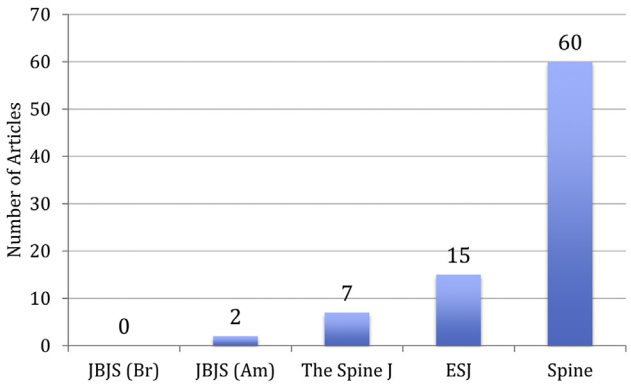

Fig. 2. Patient-reported outcome use by journal. ESJ, *European Spine Journal*; JBJS, *Journal of Bone and Joint Surgery*.

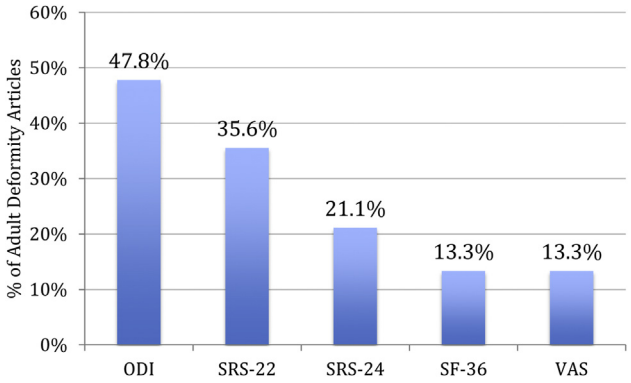

Fig. 3. Patient reported outcome prevalence. n = 84 articles. ODI, Oswestry Disability Index; SRS, Scoliosis Research Society questionnaire; SF-36, Short Form-36; VAS, visual analog scale.

ODI (43 uses; 47.8% of articles used this PRO), SRS-22 (32 uses, 35.6%), and SRS-24 (19 uses, 21.1%), and SF-36 and VAS were tied (12 uses, 13.3%) (Fig. 3). Although the ODI was the most frequently used single PRO instrument, the combined frequency of all SRS versions was significantly higher (70.2%). Excluding the top 5 PROs, all other named PROs appeared a total of 43 times.

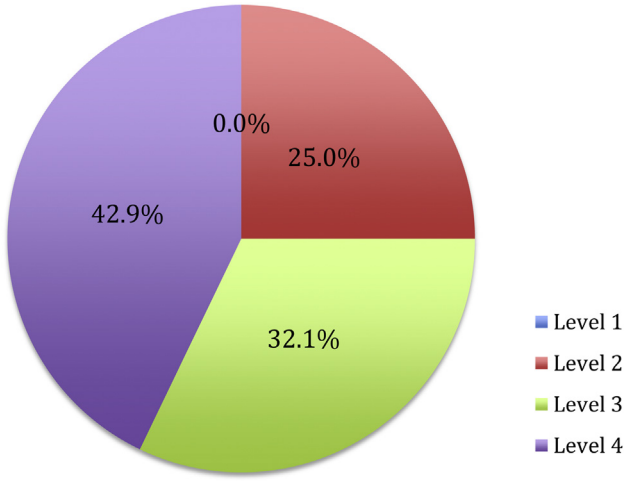

Fig. 4. Level of evidence of articles using patient-reported outcomes. n = 84 articles.

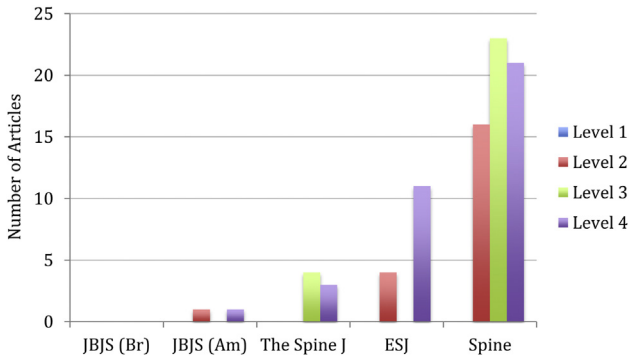

Fig. 5. Level of evidence by journal. n = 84 articles. ESJ, *European Spine Journal*; JBJS, *Journal of Bone and Joint Surgery*.

LOE 4 was most common among the articles (36 articles, 42.9%) (Fig. 4). LOE 3 was the second most common (27, 32.1%), followed by LOE 2 (21, 25.0%) and LOE 1 (0, 0.0%). Spine published the greatest number of studies having LOE 2 evidence (Fig. 5).

Errors and Inconsistencies of Reporting

The methods and results of studies reporting ODI, SRS-22, or SRS-24 scores were analyzed for errors and inconsistencies since these were the 3 most frequently used PROs. The studies are categorized in Table 2 according to their reporting inconsistency. Incomplete preoperative or postoperative data were most common across all 3 instruments, appearing in 16% of ODI studies, 58% of SRS-24 studies, and 22% of SRS-22 studies. Graphical reporting and reporting only the mean change in scores were relatively infrequent, appearing in $<10\%$ of studies for each PRO. Standard instruments were modified in 5 studies identified in this review, 3 of which combined scores from SRS-24 and SRS-22 without use of valid conversions.

Table 2
Errors and Inconsistencies of Reporting.

	Errors and inconsistencies	ODI	SRS-24	SRS-22
1	Pre- or postoperation data missing (%)	7 (16%)	11 (58%)	7 (22%)
2	Only graphical or statistical data reported (%)	2 (5%)	0 (0%)	3 (9%)
3	Only mean change reported (%)	1 (2%)	0 (0%)	1 (3%)
4	Modification of a standard instrument (%)	1 (2%)	4 (21%)	3 (9%)
5	No Error or Inconsistency (%)	32 (74%)	4 (21%)	18 (56%)
	Total	43	19	32

Discussion

The use of PRO instruments increased considerably in ASD studies over the past decade, thanks to the introduction of a validated, disease-specific PRO in the SRS-24 and an increased emphasis throughout the health care industry on measuring patient quality of life. The SRS-22, which is the product of multiple revisions of the SRS-24, and the ODI have emerged as standard PROs in ASD research. However, the frequent use of several uncommon and unnamed instruments reflects the lack of standardization in PRO instrument use in ASD research.

Comparing the findings of this study to a previous ASD literature review by Bridwell et al. [4] reveals an increase in both the frequency of PRO use and the quality of PRO studies. In a review of Spine and JBJS articles, Bridwell et al. found no prospective studies from 1996 to 2006. In our review of articles published from 2004 to 2013, we identified 22 prospective studies, with half of those published in 2011 or later. The quality of evidence in PRO studies has therefore improved since the decade studied by Bridwell et al.; however, LOE 1 studies have yet to be published.

Errors and Inconsistencies of Reporting

Errors and inconsistencies of reporting present a barrier to interstudy comparison that could be eliminated if standards of use for PROs are established. Since no standards currently exist for PRO reporting, the errors and inconsistencies evaluated in this review are a synthesis of problems with PRO use identified previously in the literature [4] and apparent barriers to interstudy comparison. The largest area of PRO reporting inconsistencies was missing preoperative or postoperative data. As Bridwell et al. [4] points out in an earlier review of ASD literature, studies without preoperative and postoperative data do not answer the question of whether ASD patients benefit from surgical treatment. Although postoperative scores are often readily obtained for retrospective studies, the postoperative scores alone are insufficient for assessing whether a patient has benefited from an intervention. Bridwell et al. identified a single study with complete preoperative and postoperative data from 1996 to 2006. This number improved considerably over the past decade, yet many articles are still missing either preoperative or postoperative data.

Exclusively graphical or statistical reporting was relatively uncommon; however, we identified studies that reported correlation statistics between PRO scores and other variables while omitting mean PRO scores. Reporting only mean change in PRO scores from preoperative to postoperative assessment was also infrequent, appearing in 1 study. Other studies reported only domain scores for a given instrument (ie, the reporting of only domain scores from SRS questionnaires or only component scores from SF-36). Domain scores are helpful in that they allow more precise comparisons to be made among outcome studies, yet studies that report certain, but not all, domains are incomplete.

Although graphical reporting and domain score reporting are mainly inconsistencies that present barriers to meta-analyses or interstudy comparisons, modifications of standard instruments are true errors of reporting. By modifying standard instruments directly (e.g. removing or replacing a question) or lumping scores from different versions of an instrument (e.g. SRS-22 and SRS-24) without using valid conversions, investigators inadvertently publish PRO data that are not valid. The modification seen most commonly among SRS-22 and SRS-24 studies was grouping and reporting scores measured from different SRS questionnaire versions. This may be attributable to the successive revisions of the SRS questionnaire over a short period. Authors and editors should be aware of these inconsistencies and errors of reporting.

Scoliosis Research Society Questionnaires

The SRS-24 was first introduced in 1999 as a disease-specific PRO instrument for adolescent scoliosis and was soon validated in the ASD population [5]. It has been sequentially revised from SRS-24 to SRS-23, then to SRS-22, and finally to SRS-22r (refined) to improve its psychometric properties [6]. The SRS-22r has 5 domains—function, pain, self-image, mental health, and satisfaction—and is the most widely validated version [6-13]. Higher scores on all SRS instruments indicate better outcomes. Because the scores of each SRS version are not directly comparable, Lai et al. developed a regression equation in 2011 to allow valid conversion from SRS-24, -23, and -22 to SRS-22r equivalents [6]. We encountered several studies that drew invalid comparisons between SRS instruments prior to and since the publication of Lai's study. These included 2 studies mentioned in the errors and inconsistencies of reporting analysis and a review paper that translated all SRS version scores into SRS-30 equivalents for

statistical purposes while acknowledging that a valid conversion method was not used. This approach is convenient but should be avoided, as it does not provide an accurate assessment of outcomes for those particular interventions investigated.

A standard minimum follow-up period may be helpful for future comparisons of PRO studies. The present analysis identified studies with follow-up periods of <1 year, 2 years, and >2 years, with a 2-year follow-up being most common. Multiple authors recommend a 2-year minimum follow-up so that the effects of chronic or subacute complications are accounted for in PROs [2, 4]. However, Glassman et al. studied and proposed a 1-year minimum follow-up standard, showing evidence that 1-year outcomes are predictive of 2-year outcomes [14]. Results from prospective studies with high quality evidence would be available sooner with this method. Given the disagreement in the literature as well as the potential benefits of a shorter minimum follow-up, consensus regarding an optimal follow-up period for ASD outcome studies is needed.

ODI

The ODI was introduced in 1980 to measure clinical outcomes in back pain research [15]. It has been referred to as a regional-specific, rather than disease-specific, PRO instrument in ASD [1] because it addresses only low back pain concerns rather than the full spectrum of quality-of-life concerns that patients with ASD face. The most common presenting symptoms of ASD are neurologic dysfunction and pain, so this instrument addresses 1 of the key symptoms [16]. In general, the ODI is a short, easy-to-use PRO instrument with only 10 questions, making it a favorite for clinical outcomes research in ASD. Despite widespread use of the ODI, good validation studies of the instrument's psychometric properties are lacking in the ASD population. In a point of view letter to Spine, Walsh points out that wide use of an instrument should not be mistaken for validation of an instrument [17-19]. Though the ODI is shown in this review to be used widely in ASD, it should not be considered a gold-standard PRO until its reliability (internal consistency and reproducibility) and validity (criterion, concurrent, and discriminant validity) have been studied in the ASD population [8, 17].

Like the SRS, the ODI has undergone multiple revisions since its introduction. Fairbank, the original author of the ODI, recognizes versions v1.0 (1980) [15], v2.0 (2000) [20], v2.1 (2000) [21], and v2.1a of the ODI instrument. The version of ODI used in a given study, however, is rarely reported. This may be due to users' lack of familiarity with the history and development of the instrument. Poor version reporting makes comparisons of ODI scores between studies problematic as scores from different versions are not directly comparable. In a 2007 review paper titled "The Use and Abuse of the Oswestry Disability Index," Fairbank et al. made an effort to break the cycle of poor version reporting, stating that, "it is essential that investigators indicate the version number of an instrument in publications and that this practice is enforced by editors and referees" [22]. Meta-analyses and systematic reviews of ASD literature should also acknowledge the lack of ODI version information to help reverse the trend of poor version reporting.

Short Form-36

The SF-36 is a general-health questionnaire that enables comparisons of quality of life outcomes across medical disciplines. The instrument was first published in 1992 by Ware et al. [23] and has been validated in several spinal conditions, including back pain, spinal injury and disk herniation [24,25]. SF-36 is the only PRO instrument among the top 5 that is routinely used to calculate health utility scores, which are the core of cost-effectiveness studies. Health utility scores on a scale of 0 to 1 are considered over time to determine quality-adjusted life-years (QALYs) resulting from a given treatment. Costs per QALY are then compared in cost-effectiveness studies. Our analysis showed that SF-36 was used in only 13.3% of outcomes studies in ASD over the past decade. Considering the 16-fold estimated increase in ASD spending over the past decade, general-health questionnaires like SF-36, SF-12, and EQ-5D will need to be used more frequently to support cost-effectiveness research prompted by this rise in spending [1].

Regression equations have been developed to conveniently convert ODI to health utility scores in the lumbar degenerative disease population [26,27]. The benefit of this approach is a reduced administrative burden for researchers; however, the population for which the regression equations were developed is not specifically spinal deformity, making it less than ideal for ASD studies.

Visual Analog Scale

The VAS has been used in clinics and research for measuring pain since the 1920s [28]. It is primarily used to measure back and leg pain, but has been employed to measure a diversity of other endpoints such as satisfaction with treatment and even health utility scores [29,30]. A common VAS layout is a 100-mm line with the text "no pain" on the left and "severe pain" on the right. Patients are instructed to pick the location along the line that best describes their condition. Many variations of the orientation and text of the VAS scale exist. Scores are frequently reported from 0 to 10, and sometimes from 0 to 100. The VAS is a very quick and convenient PRO that assesses one of the most common symptoms of ASD—pain. However, the VAS cannot be used for cost-effectiveness studies and only provides a tangential and limited assessment of patient outcome following surgical intervention.

As with any study, ours is not without shortcomings. The main one is that we did not include all pertinent spine-related journals, including neurosurgical journals, in our search. It would have been too labor intensive, perhaps logistically impossible, to include all journals. We modeled

our study design after that of Hunt et al. [31] and chose the 5 journals based on high impact factor and broad readership for select sampling. In reviewing candidate articles for our analysis, it is possible that PRO studies were missed, especially if no mention was made of PRO use in the title or abstract. This is the first study of adult spinal deformity research to measure the frequency of use of PROs and errors and inconsistencies of reporting PRO scores. Our results afford the insights needed to establish standards of PRO use going forward.

Conclusion

In this review, we have shown an increase in PRO studies in ASD surgery research over the past decade and quantified the frequency of the most used instruments. For the growing body of data generated by PRO studies to be useful and reliable going forward, (1) the 24 unique PROs used in ASD must be pared down to a few PROs that can be used consistently, and (2) authors must take care to accurately report instrument version and (3) abstain from modifying standard instruments. There is a need to standardize PRO use in ASD research, and our study hopefully provides background data for that purpose.

References

[1] McCarthy I, Hostin R, O'Brien M, et al. Health economic analysis of adult deformity surgery. *Neurosurg Clin N Am* 2013;24:293–304.

[2] Yadla S, Maltenfort MG, Ratliff JK, et al. Adult scoliosis surgery outcomes: a systematic review. *Neurosurg Focus* 2010;28:E3.

[3] OCEBM Levels of Evidence Working Group. Levels of Evidence (March 2009) [Oxford Center for Evidence-Based Medicine web site]. http://www.cebm.net/ocebm-levels-of-evidence/ Accessed June 6, 2014.

[4] Bridwell KH, Berven S, Edwards 2nd C, et al. The problems and limitations of applying evidence-based medicine to primary surgical treatment of adult spinal deformity. *Spine (Phila Pa 1976)* 2007;32(19 Suppl):S135–9.

[5] Berven S, Deviren V, Demir-Deviren S, et al. Studies in the modified Scoliosis Research Society Outcomes Instrument in adults: validation, reliability, and discriminatory capacity. *Spine (Phila Pa 1976)* 2003;28:2164–9; discussion 2169.

[6] Lai SM, Burton DC, Asher MA, et al. Converting SRS-24, SRS-23, and SRS-22 to SRS-22r: establishing conversion equations using regression modeling. *Spine (Phila Pa 1976)* 2011;36:E1525–33.

[7] Alanay A, Cil A, Berk H, et al. Reliability and validity of adapted Turkish Version of Scoliosis Research Society-22 (SRS-22) questionnaire. *Spine (Phila Pa 1976)* 2005;30:2464–8.

[8] Asher M, Min Lai S, Burton D, et al. The reliability and concurrent validity of the Scoliosis Research Society-22 patient questionnaire for idiopathic scoliosis. *Spine (Phila Pa 1976)* 2003;28:63–9.

[9] Burton DC, Glattes RC. Measuring outcomes in spinal deformity. *Neurosurg Clin N Am* 2007;18:403–5.

[10] Climent JM, Bago J, Ey A, et al. Validity of the Spanish version of the Scoliosis Research Society-22 (SRS-22) Patient Questionnaire. *Spine (Phila Pa 1976)* 2005;30:705–9.

[11] Glattes RC, Burton DC, Lai SM, et al. The reliability and concurrent validity of the Scoliosis Research Society-22r patient questionnaire compared with the Child Health Questionnaire-CF87 patient questionnaire for adolescent spinal deformity. *Spine (Phila Pa 1976)* 2007;32:1778–84.

[12] Hashimoto H, Sase T, Arai Y, et al. Validation of a Japanese version of the Scoliosis Research Society-22 Patient Questionnaire among idiopathic scoliosis patients in Japan. *Spine (Phila Pa 1976)* 2007;32:E141–6.

[13] Mousavi SJ, Mobini B, Mehdian H, et al. Reliability and validity of the Persian version of the Scoliosis Research Society-22r questionnaire. *Spine (Phila Pa 1976)* 2010;35:784–9.

[14] Glassman SD, Schwab F, Bridwell KH, et al. Do 1-year outcomes predict 2-year outcomes for adult deformity surgery? *Spine J* 2009;9:317–22.

[15] Fairbank JC, Couper J, Davies JB, et al. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980;66:271–3.

[16] Heary RF. Evaluation and treatment of adult spinal deformity. Invited submission from the Joint Section Meeting on Disorders of the Spine and Peripheral Nerves. *J Neurosurg Spine* 2004;2004:9–18.

[17] Ware Jr JE. Standards for validating health measures: definition and content. *J Chronic Dis* 1987;40:473–80.

[18] Walsh T. Point of view on the Oswestry Disability Index. *Spine (Phila Pa 1976)* 2000;25:2953.

[19] Deyo RA, Patrick DL. Barriers to the use of health status measures in clinical investigation, patient care, and policy research. *Med Care* 1989;27(3 suppl):S254–68.

[20] Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine (Phila Pa 1976)* 2000;25:2940–52; discussion 2952.

[21] Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine (Phila Pa 1976)* 2000;25:3115–24.

[22] Fairbank JC. Use and abuse of Oswestry Disability Index. *Spine (Phila Pa 1976)* 2007;32:2787–9.

[23] Ware Jr JE, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–83.

[24] Ware J. SF-36 Health Survey Update [SF-36 web site]. October 2, 2003. Available at: http://www.sf-36.org/tools/sf36.shtml. Accessed June 6, 2014.

[25] Guilfoyle MR, Seeley H, Laing RJ. The Short Form 36 health survey in spine disease—validation against condition-specific measures. *Br J Neurosurg* 2009;23:401–5.

[26] Carreon LY, Bratcher KR, Das N, et al. Estimating EQ-5D Values From the Oswestry Disability Index and Numeric Rating Scales for Back and Leg Pain. *Spine (Phila Pa 1976)* 2014;39:678–82.

[27] Carreon LY, Glassman SD, McDonough CM, et al. Predicting SF-6D utility scores from the Oswestry disability index and numeric rating scales for back and leg pain. *Spine (Phila Pa 1976)* 2009;34:2085–9.

[28] Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. *Res Nurs Health* 1990;13:227–36.

[29] Stiggelbout AM, Eijkemans MJ, Kiebert GM, et al. The "utility" of the visual analog scale in medical decision making and technology assessment. Is it an alternative to the time trade-off? *Int J Technol Assess Health Care* 1996;12:291–8.

[30] van den Akker ME, Arts MP, van den Hout WB, et al. Tubular diskectomy vs conventional microdiskectomy for the treatment of lumbar disk-related sciatica: cost utility analysis alongside a double-blind randomized controlled trial. *Neurosurgery* 2011;69:829–35; discussion 835–6.

[31] Hunt KJ, Hurwit D. Use of patient-reported outcome measures in foot and ankle research. *J Bone Joint Surg Am* 2013;95:e118 (1–9).