

Current Trends in the Use of Patient-Reported Outcome Instruments in Degenerative Cervical Spine Surgery

Haruki Ueda¹ Holt S. Cutler¹ Javier Z. Guzman¹ Samuel K. Cho¹

¹Department of Orthopaedic Surgery, Icahn School of Medicine at Mount Sinai, New York, New York, United States

Global Spine J 2016;6:242–247.

Address for correspondence Samuel K. Cho, MD, Department of Orthopaedic Surgery, Icahn School of Medicine at Mount Sinai, 5 East 98th Street, Box 1188, New York, NY 10029, United States (e-mail: samuel.cho@mountsinai.org).

Abstract

Study Design Bibliometric analysis.

Objective To determine trends, frequency, and distribution of patient-reported outcome instruments (PROIs) in degenerative cervical spine surgery literature over the past decade.

Methods A search was conducted via PubMed from 2004 to 2013 on five journals (*The Journal of Bone and Joint Surgery*, *The Bone and Joint Journal*, *The Spine Journal*, *European Spine Journal*, and *Spine*), which were chosen based on their impact factors and authors' consensus. All abstracts were screened and articles addressing degenerative cervical spine surgery using PROIs were included. Articles were then analyzed for publication date, study design, journal, level of evidence, and PROI trends. Prevalence of PROIs and level of evidence of included articles were analyzed.

Results From 19,736 articles published, 241 articles fulfilled our study criteria. Overall, 53 distinct PROIs appeared. The top seven most frequently used PROIs were: Japanese Orthopaedic Association score (104 studies), visual analog scale for pain (100), Neck Disability Index (72), Short Form-36 (38), Nurick score (25), Odom criteria (21), and Oswestry Disability Index (15). Only 11 PROIs were used in 5 or more articles. Thirty-three of the PROIs were appeared in only 1 article. Among the included articles, 16% were of level 1 evidence and 32% were of level 4 evidence.

Conclusion Numerous PROIs are currently used in degenerative cervical spine surgery. A consensus on which instruments to use for a given diagnosis or procedure is lacking and may be necessary for better communication and comparison, as well as for the accumulation and analysis of vast clinical data across multiple studies.

Keywords

- ▶ patient-reported outcome instrument
- ▶ degenerative cervical spine
- ▶ cervical spine surgery
- ▶ clinical scales
- ▶ cost effectiveness

Introduction

Patient-centered health measures have been recognized as valuable and useful tools to report and quantify the efficacy of surgical procedures.^{1–3} Patient-centered health measures also enable outcome comparisons of cost-effectiveness across studies. Currently, there are many kinds of patient-reported outcome instruments (PROIs) in the field of cervical spine

surgery without an established consensus as to which instrument should be used for a particular diagnosis or procedure.⁴ Without a consensus, long-term efficacy or cost-effectiveness of different treatments becomes difficult to assess.

The purpose of this study was to determine the frequency and distribution of various PROIs used in the degenerative cervical spine surgery literature and to identify the trends of use of these instruments over the past decade.

received
April 14, 2015
accepted after revision
June 22, 2015
published online
July 28, 2015

DOI <http://dx.doi.org/10.1055/s-0035-1559584>.
ISSN 2192-5682.

© 2016 Georg Thieme Verlag KG
Stuttgart · New York

License terms



Materials and Methods

A search was conducted on PubMed for publications from 2004 to 2013 in five major orthopedic and spine journals—*The Journal of Bone and Joint Surgery*, *The Bone and Joint Journal* (formerly *The Journal of Bone and Joint Surgery [British Volume]*), *The Spine Journal*, *European Spine Journal*, and *Spine*—which were chosen based on the authors' consensus as high-impact journals with respected quality standards. Articles with a clinical focus on cervical spine surgery for degenerative cervical spine diseases, including cervical spondylotic myelopathy, cervical radiculopathy, ossification of posterior longitudinal ligament, cervical spondylosis, cervical degenerative disk disease, and cervical disk hernia, were identified and chosen for abstract review according to title relevance. All the abstracts were screened, and articles that addressed degenerative cervical spine surgery and used PROIs were included. Review articles were excluded.

All the included articles were then thoroughly reviewed and analyzed for their journal, publication date, study design, level of evidence (LOE), and preoperative and 2-year follow-up scores for the most commonly used instruments. The LOE was assigned based on the definitions by the Oxford Center for Evidence-Based Medicine, and only articles with LOEs of 1 to 4 were included.⁵ This study did not need Institutional Review Board approval.

Results

From 19,736 articles published,⁵ we identified 1,079 articles reporting the use of PROIs in the field of spine surgery. Of these articles, 241 met our study criteria and were included (► **Table 1**). Articles published per journal in descending order were *Spine* (40%), *European Spine Journal* (35%), *The Spine Journal* (17%), *The Bone and Joint Journal* (4%), and *The Journal of Bone and Joint Surgery* (4%; ► **Fig. 1**). Overall, there were 53 distinct outcome measures utilized.

An upward trend in PROI usage was observed over the past decade (► **Fig. 2**). Of note, about one third of these articles (30.7%, 74 articles) were published in the previous 2 years (2012 to 2013).

The five most used outcome measures in descending order were: the Japanese Orthopaedic Association (JOA) score (104 studies, 43.2%), visual analog scale (VAS) for pain (100 studies,

Table 1 The number of studies that contained patient-reported outcome instruments among the five journals studied

| Journal | No. of studies |
|--|----------------|
| <i>European Spine Journal</i> | 85 |
| <i>The Journal of Bone and Joint Surgery</i> | 8 |
| <i>The Bone and Joint Journal</i> | 10 |
| <i>Spine</i> | 97 |
| <i>The Spine Journal</i> | 41 |
| Total | 241 |

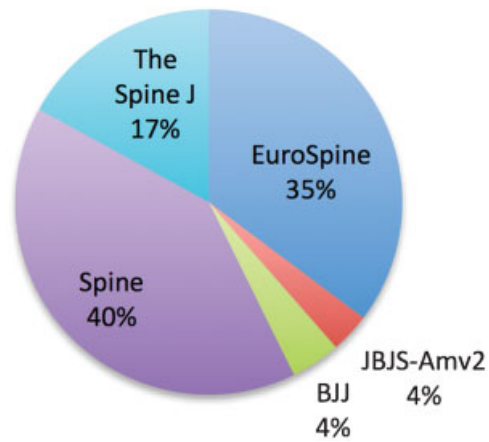


Fig. 1 The contribution of each journal to the total of included articles. Abbreviations: EuroSpine, *European Spine Journal*; JBJS-Amv2, *Journal of Bone and Joint Surgery American Volume*; BJJ, *The Bone and Joint Journal*; The Spine J, *The Spine Journal*.

41.5%), Neck Disability Index (NDI; 72 studies, 29.9%), Short Form-36 (SF-36; 38 studies, 15.8%), and Nurick score (25 studies, 10.4%; ► **Fig. 3**).⁶⁻¹² Of the 53 total PROIs, only 11 were used in 5 or more articles (► **Table 2**). Thirty-three of the PROIs appeared in only a single article. One third of the articles were of LOE 4 (32.0%), and only 16.4% of all articles were of LOE 1 (► **Fig. 4**).

Discussion

With the advent of evidence-based medicine, research in the field of spine surgery increasingly requires objective statistical analyses of the treatment outcomes. Accordingly, researchers have responded by turning to PROIs to evaluate the health care outcomes.¹⁻³ PROIs are powerful tools for evaluating clinical results and objectively quantifying outcome success. In this study, the use of PROIs in degenerative cervical spine surgery was found to be increasing in five top-impact orthopedic and spine surgery journals in the past 10 years. Among the PROIs used in these journals, ~30% occurred in the previous 2 years (2012 to 2013) and ~41% in the previous 3 years (2011 to 2013).⁶⁻¹² Among the 241 articles, PROIs were used 465 times, a mean of 1.9 PROIs per study. Among the top PROIs, the JOA and NDI were used the most over the past decade. Although increasing, the SF-36 was not utilized significantly over the past decade despite the value of this health-related quality-of-life instrument in calculating quality-adjusted life-years and analyzing cost-effectiveness. Given that cost-effectiveness studies will be of increasing interest in the coming years, investigators should consider using the SF-36 or other validated scales with the capability of providing cost-effectiveness data (► **Fig. 5**).

There are many different PROIs available, which often makes the choice of which to use unclear.^{4,13,14} In fact, in the past 10 years of degenerative cervical spine surgery studies, 53 unique PROIs were used. Many of these PROIs (33) were used only once. Two factors are mainly responsible

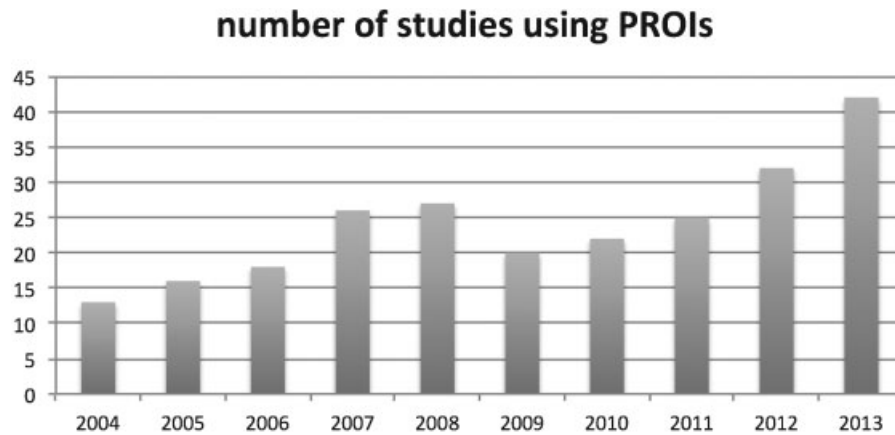


Fig. 2 Overall trend of the use of patient-reported outcome instruments (PROIs) from 2004 to 2013 in the field of degenerative cervical spine surgery.

for the diversity of PROIs used. First, cervical spine disease can present with various clinical symptoms ranging from localized neck pain to systemic myelopathy, and therefore a single scoring system may not be sufficient to address all cervical spine outcomes. Second, no recommendation or clear guideline exists for the selection of PROIs in cervical spine studies, leaving investigators to choose from numerous existing PROIs or design their own custom tool. Although PROI use is increasing overall, the wide range of PROIs utilized complicates comparisons that we would hope to draw between studies.

Outcome measures are classified into three general categories: general quality of life, pain, and disease-specific outcomes.⁸ Hunt and Hurwit recommended the use of combined disease-specific and general health PROIs.¹⁵ For example, among the top seven PROIs, the SF-36 is a general health measure, the VAS is a pain scale, and the JOA, NDI, Nurick, and ODI are disease-specific measures. As a clinical scoring system that can be utilized widely, the ideal PROI must be valid, reliable, and responsive.^{15,16} Multiple studies validated these PROIs by demonstrating a correlation with other PROIs—SF-

36 in comparison with MDI and VAS⁷; SF-36 with VAS, NDI, and MDI¹⁷; SF-36 with VAS and Nurick^{18,19}; SF-12 with SF-36²⁰; Nurick with modified JOA and JOA.^{10,21,22} In addition, interobserver and test-retest reliability of these most frequently used PROIs have been surveyed including for JOA,²³ NDI and VAS,²⁴ SF-36,^{17,25} and SF-12.²⁶

Disease-Specific Outcome Measures

As mentioned previously, cervical spine pathology can present with diverse clinical symptoms and requires the appropriate choice of PROI for each diagnosis. Kalsi-Ryan et al recently published a statement regarding the use of PROIs for cervical spondylotic myelopathy.²⁷ They recommended that reliable, valid, and responsive PROIs be used, as well as ancillary measures to record and evaluate patients' function. In particular, the authors recommended that the modified JOA, Nurick grade, NDI, Myelopathy Disability Index (MDI), and 30 m Walk Test should be used with the ancillary measures such as the shortened Disabilities of the Arm, Shoulders and Hand questionnaire; Berg Balance Scale; Graded Redefined of Strength Sensitivity and Prehension; Grip Strength; and GAITRite analysis.²⁷ Among the disease-specific PROIs, we recommend the JOA scale because it provides comprehensive coverage of myelopathic symptoms including bowel and bladder functions, as opposed to the Nurick, NDI, MDI, and 30 m Walk Test, which only assess upper or lower extremity function. The JOA score was also the most frequently used PROI in this study.

When it comes to cervical radiculopathy, however, the choice may be different from those for cervical myelopathy given the difference in symptoms. The questionnaires should contain items that can distinguish radicular arm pain from axial neck pain. From this perspective, the Nurick score that classifies radicular-only symptom as grade 0 will not be suitable nor sufficient to describe this group of patients well. Although the modified JOA and the European Myelopathy Scale are widely accepted as reliable and responsive instruments, they are designed to assess long tract signs including lower extremity, bowel, and bladder functions more than upper extremity symptoms and thus may not

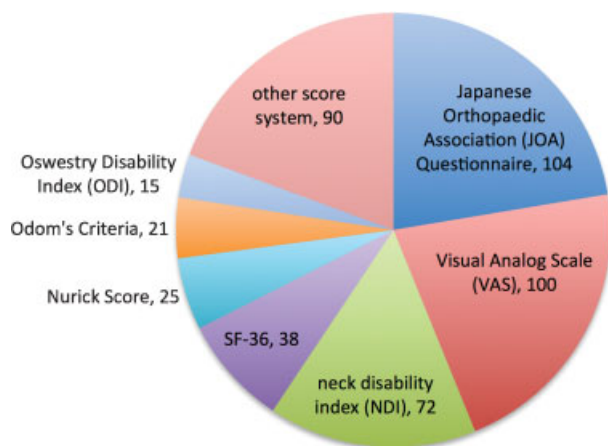


Fig. 3 Frequency of appearance of each patient-reported outcome instrument in articles from 2004 to 2013.

Table 2 Most commonly used PROIs

| PROI | No. of articles | % of total articles |
|----------------------|-----------------|---------------------|
| JOA score | 104 | 43.2 |
| VAS | 100 | 41.5 |
| NDI | 72 | 29.9 |
| F-36 | 38 | 15.8 |
| Nurick score | 25 | 10.4 |
| Odom's criteria | 21 | 8.7 |
| ODI | 15 | 6.2 |
| Modified JOA | 12 | 5.0 |
| Global Outcome Scale | 9 | 3.7 |
| EQ-5D | 7 | 2.9 |
| SF-12 | 6 | 2.5 |
| CSOQ | 5 | 2.1 |

Abbreviations: EQ-5D, Euro-QoL-5D; JOA, Japanese Orthopaedic Association; NDI, Neck Disability Index; ODI, Oswestry Disability Index; PROI, patient-reported outcome instrument; SF-12, Short Form-12; SF-36, Short Form-36; SOQ, Cervical Spine Outcomes Questionnaire; VAS, visual analog scale.

Note: Studies used more than 4 times with the number of appearance and % of appearance among total 241 studies. Besides these PROIs, 2 scoring systems were used 3 times, 6 systems were used twice, and 33 systems were used only once. In total, 465 PROIs appeared in 241 studies.

accurately reflect the intensity or change of cervical radiculopathy.

Cleland et al and Young et al studied the usefulness of the NDI, Patient-Specific Functional scale, and Numeric Pain Rating Scales in patients with radiculopathy.^{24,28} Although they came to different conclusions regarding the test-retest reliability of the Patient-Specific Functional scale, these investigators thought there was fair reliability for both the NDI and Numeric Pain Rating Scales, and adequate responsiveness for all three scales. Among these scales, the NDI is currently

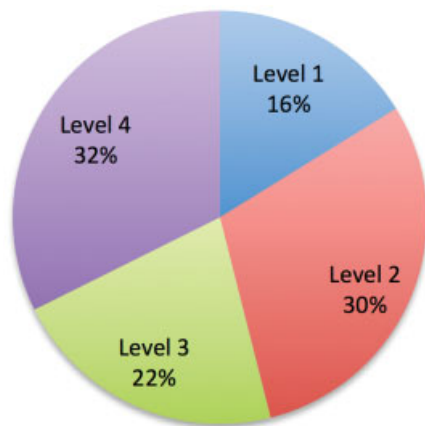


Fig. 4 Distribution of the level of evidence.

widely used, has been proven valid and reliable in multiple studies, and thus may be an ideal choice for cervical radiculopathy.^{29,30}

General Health Questionnaires

Among the most frequent PROIs in this study, the general health questionnaires were the Euro-QoL-5D (EQ-5D), SF-36, and Odom criteria. The Odom criteria focus on the patients' recovery of their preoperative symptoms following surgical intervention and do not truly assess their general health.¹² In our analysis, we observed a sharp decline in its use, and the Odom criteria may eventually be replaced by another general health PROI such as the EQ-5D.

The EQ-5D is a relatively new five-item scale that assesses health in terms of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Its brevity and simplicity are the main advantages. Use of the EQ-5D is currently on the rise, and we may see more of this PROI in the future.

The SF-36 is older than the EQ-5D and is another well-established questionnaire for general health. It consists of 36 items. The SF-36 evaluates functional status (19 items), well-being (11 items), general health perception (5 items), and health change over the past year (1 item). This questionnaire is comprehensive and allows for a sensitive and detailed analysis, but compared with the EQ-5D, it is significantly more complex and time-consuming.

Given the above, the EQ-5D may be a reasonable choice to measure generic health as it covers the most important health-related quality-of-life dimensions for musculoskeletal diseases in five simple questions.³¹ However, one weakness of the EQ-5D is the ceiling effect due to its relative simplicity. In one study, 32% of patients reached highest or lowest values at follow-up.³² The large ceiling effect for the EQ-5D suggests that it would be more suitable for measuring the health of populations with more morbidity, where the distribution of the total score is less skewed.³³ For example, elderly responders tend to have multiple morbidities and more difficulties with answering complicated questions. The EQ-5D may be more appropriate when researchers target this kind of patient group.

To simplify the SF-36 without substantial loss of information, the eight-scale profile was reduced to two summary measures of health status, creating the SF-12.³³ The SF-12 contains questions regarding physical functioning, role-physical, role-emotional, mental health, bodily pain, vitality, social functioning, and general health. The former four have two items each, and the latter four have one item each. Jenkinson et al compared the SF-12 and SF-36 scores in various patient groups and concluded that these two instruments indicated the same magnitude of ill health and degree of change over time and that wherever two summary scores of health status are adequate, the SF-12 may be the instrument of choice.³⁴ For surveying musculoskeletal diseases including spinal pathologies, the SF-12 may be adequate. In addition, a comparison study by Johnson and Coons showed that the SF-12 was more sensitive than the EQ-5D. They found that for the

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

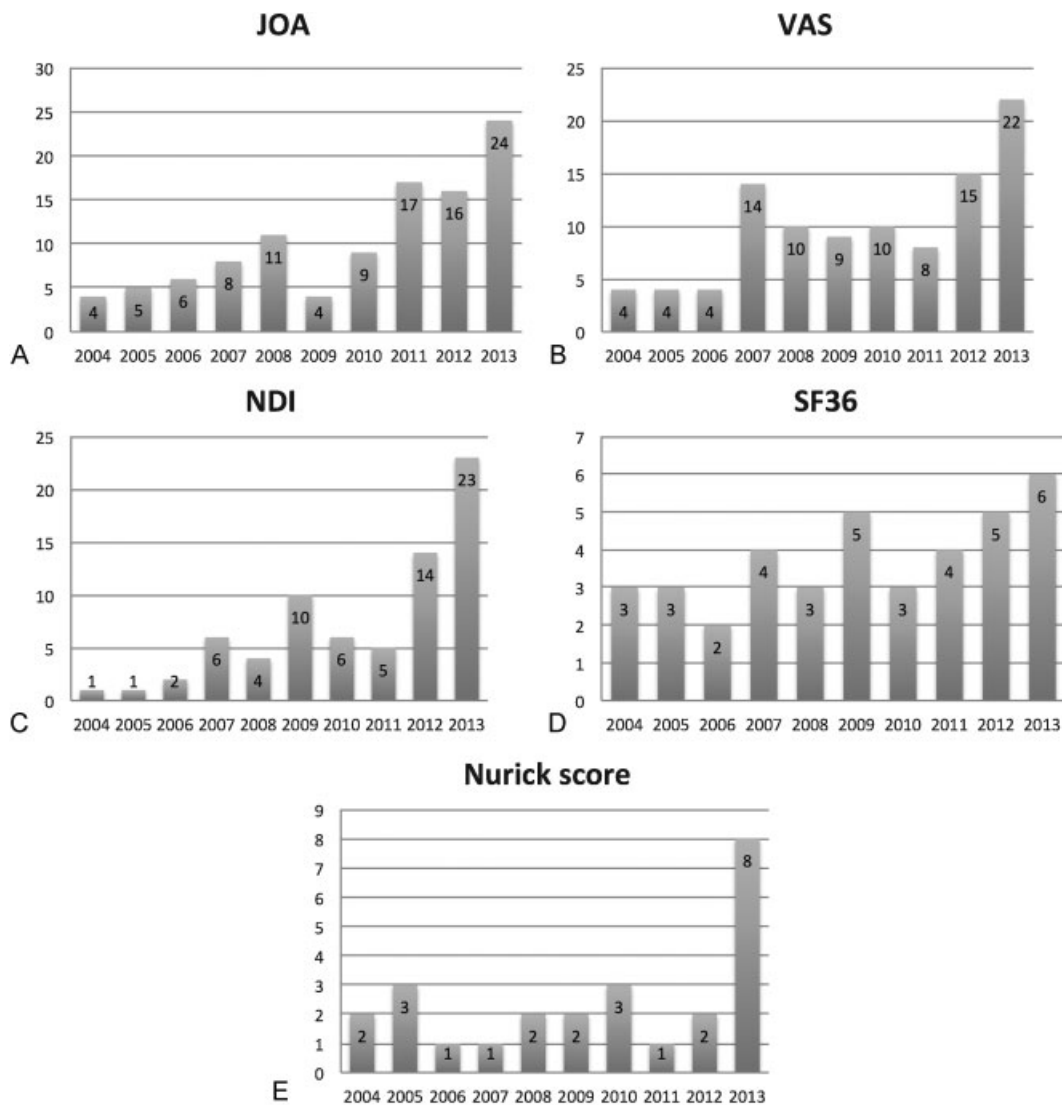


Fig. 5 Trend of the use of each top-five patient-reported outcome instruments: (A) (JOA), (B) visual analog scale (VAS), (C) Neck Disability Index (NDI), (D) Short Form-36 (SF36), (E) Nurick score.

subjects who reported no problems on the EQ-5D dimensions, the SF-12 component scores were still able to differentiate people with and without specific chronic medical conditions.³³ We believe either SF-12 or EQ-5D can be employed for general health surveys for degenerative cervical spinal pathologies with the understanding of its benefits and limitations.

Our study has a few limitations that the authors would like to acknowledge. The main one is that we did not include neurosurgical journals in our review. However, *Spine* and *The Spine Journal* both have a neurosurgical audience and therefore capture this perspective. A second limitation is that relevant studies within the five journals we reviewed may have been missed during screening if PROIs or outcome measurements were not mentioned in the title or abstract.

Conclusions

The breadth of PROIs in degenerative cervical spine surgery research is extensive. A consensus on which instruments should be used for a given diagnosis or procedure may be necessary for better communication and comparison, as well as for the accumulation and analysis of vast clinical data across multiple studies. Additionally, validity and reliability studies for several of the top PROIs are incomplete and warrant further research as their PROI usage continues to increase. For degenerative cervical spine surgery, we recommend a combination of general health and cervical spine disease-specific PROIs, along with additional scales for pain (e.g., VAS) on an as-needed basis. Moreover, a consensus on PROI use in cervical spine surgery may be necessary to improve consistency and allow for interstudy comparisons in the future.

Disclosures

Haruki Ueda, none
 Holt S. Cutler, none
 Javier Z. Guzman, none
 Samuel K. Cho, Consultant: Stryker

Acknowledgment

This study was sponsored by the AOSpine Research Commission Mentoring Program.

References

- Rihn JA, Berven S, Allen T, et al. Defining value in spine care. *Am J Med Qual* 2009;24(6, Suppl):4S-14S
- Rihn JA, Currier BL, Phillips FM, Glassman SD, Albert TJ. Defining the value of spine care. *J Am Acad Orthop Surg* 2013;21(7):419-426
- Porter ME. What is value in health care? *N Engl J Med* 2010;363(26):2477-2481
- Greenough CG. Outcome assessment: recommendations for daily practice. *Eur Spine J* 2006;15(Suppl 1):S118-S123
- Centre for Evidence-Based Medicine. Oxford Centre for Evidence-based Medicine—Levels of Evidence (March 2009). Available at: <http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/>. Accessed July 17, 2015
- Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine (Phila Pa 1976)* 2000;25(22):2940-2952, discussion 2952
- 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(4):401-405
- McCormick JD, Werner BC, Shimer AL. Patient-reported outcome measures in spine surgery. *J Am Acad Orthop Surg* 2013;21(2):99-107
- Nurick S. The pathogenesis of the spinal cord disorder associated with cervical spondylosis. *Brain* 1972;95(1):87-100
- Revanappa KK, Rajshekhar V. Comparison of Nurick grading system and modified Japanese Orthopaedic Association scoring system in evaluation of patients with cervical spondylotic myelopathy. *Eur Spine J* 2011;20(9):1545-1551
- Ware JE Jr. SF-36 health survey update. *Spine (Phila Pa 1976)* 2000;25(24):3130-3139
- Zoëga B, Kärrholm J, Lind B. Outcome scores in degenerative cervical disc surgery. *Eur Spine J* 2000;9(2):137-143
- Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders. Introduction. *Spine (Phila Pa 1976)* 2000;25(24):3097-3099
- Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. *Spine (Phila Pa 1976)* 2000;25(24):3100-3103
- Hunt KJ, Hurwit D. Use of patient-reported outcome measures in foot and ankle research. *J Bone Joint Surg Am* 2013;95(16):e118, 1-9
- Deyo RA, Battie M, Beurskens AJ, et al. Outcome measures for low back pain research. A proposal for standardized use. *Spine (Phila Pa 1976)* 1998;23(18):2003-2013
- Latimer M, Haden N, Seeley HM, Laing RJ. Measurement of outcome in patients with cervical spondylotic myelopathy treated surgically. *Br J Neurosurg* 2002;16(6):545-549
- King JT Jr, Roberts MS. Validity and reliability of the Short Form-36 in cervical spondylotic myelopathy. *J Neurosurg* 2002;97(2, Suppl):180-185
- King JT Jr, Tsevat J, Moosy JJ, Roberts MS. Preference-based quality of life measurement in patients with cervical spondylotic myelopathy. *Spine (Phila Pa 1976)* 2004;29(11):1271-1280
- Singh A, Gnanalingham K, Casey A, Crockard A. Quality of life assessment using the Short Form-12 (SF-12) questionnaire in patients with cervical spondylotic myelopathy: comparison with SF-36. *Spine (Phila Pa 1976)* 2006;31(6):639-643
- Vitzthum HE, Dalitz K. Analysis of five specific scores for cervical spondylogenic myelopathy. *Eur Spine J* 2007;16(12):2096-2103
- Singh A, Crockard HA. Comparison of seven different scales used to quantify severity of cervical spondylotic myelopathy and post-operative improvement. *J Outcome Meas* 2001-2002;5(1):798-818
- Yonenobu K, Abumi K, Nagata K, Taketomi E, Ueyama K. Interobserver and intraobserver reliability of the Japanese Orthopaedic Association scoring system for evaluation of cervical compression myelopathy. *Spine (Phila Pa 1976)* 2001;26(17):1890-1894, discussion 1895
- Young IA, Cleland JA, Michener LA, Brown C. Reliability, construct validity, and responsiveness of the neck disability index, patient-specific functional scale, and numeric pain rating scale in patients with cervical radiculopathy. *Am J Phys Med Rehabil* 2010;89(10):831-839
- Brazier JE, Harper R, Jones NM, et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ* 1992;305(6846):160-164
- Luo X, George ML, Kakouras I, et al. Reliability, validity, and responsiveness of the short form 12-item survey (SF-12) in patients with back pain. *Spine (Phila Pa 1976)* 2003;28(15):1739-1745
- Kalsi-Ryan S, Singh A, Massicotte EM, et al. Ancillary outcome measures for assessment of individuals with cervical spondylotic myelopathy. *Spine (Phila Pa 1976)* 2013;38(22, Suppl 1):S111-S122
- Cleland JA, Fritz JM, Whitman JM, Palmer JA. The reliability and construct validity of the Neck Disability Index and patient specific functional scale in patients with cervical radiculopathy. *Spine (Phila Pa 1976)* 2006;31(5):598-602
- MacDermid JC, Walton DM, Avery S, et al. Measurement properties of the neck disability index: a systematic review. *J Orthop Sports Phys Ther* 2009;39(5):400-417
- McCarthy MJ, Grevitt MP, Silcocks P, Hobbs G. The reliability of the Vernon and Mior neck disability index, and its validity compared with the short form-36 health survey questionnaire. *Eur Spine J* 2007;16(12):2111-2117
- Picavet HS, Hoeymans N. Health related quality of life in multiple musculoskeletal diseases: SF-36 and EQ-5D in the DMC3 study. *Ann Rheum Dis* 2004;63(6):723-729
- Fankhauser CD, Mutter U, Aghayev E, Mannion AF. Validity and responsiveness of the Core Outcome Measures Index (COMI) for the neck. *Eur Spine J* 2012;21(1):101-114
- Johnson JA, Coons SJ. Comparison of the EQ-5D and SF-12 in an adult US sample. *Qual Life Res* 1998;7(2):155-166
- Jenkinson C, Layte R, Jenkinson D, et al. A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? *J Public Health Med* 1997;19(2):179-186