



# Lumbar Spinal Stenosis: Objective Measurement Scales and Ambulatory Status

Hussam Abou-Al-Shaar<sup>1</sup>, Owoicho Adogwa<sup>2</sup>, Ankit I. Mehta<sup>2</sup>

<sup>1</sup>Department of Neurosurgery, Hofstra Northwell School of Medicine, Manhasset, NY, USA

<sup>2</sup>Department of Neurosurgery, University of Illinois at Chicago, Chicago, IL, USA

Lumbar spinal stenosis (LSS) is one of the most common affecting the elderly population that may lead to loss of function and the inability to execute basic activities of daily living. While surgical decompression remains the standard of care, choosing an optimal management strategy is usually guided by a set of clinical, radiological, and measurement indices. However, to date, there is a major uncertainty and discrepancy regarding the methodology used. There is also inconsistent adoption of outcome measures across studies, which may result in huge limitations in predicting the efficacy and cost-effectiveness of different treatment paradigms. Herein, we review the various measurement indices used for outcome assessment among patients with LSS, and delineate the major advantages and disadvantages of each index. We call for the development of a single objective outcome measure that encompasses and addresses all issues encountered in this heterogeneous group of patients, including monitoring the patient's progression after treatment.

**Keywords:** Spinal stenosis; Outcomes; Ambulation; Surgery; Laminectomy

## Introduction

Lumbar spinal stenosis (LSS) is a frequently encountered condition in elderly patients (age >65 years) [1]. Neurogenic claudication in elderly patients may cause loss of function and an associated inability to perform basic activities of daily living. Elderly patients demand medical care, including surgery, not only to save their lives but also to increase its quality. Most patients are managed non-operatively; however, a subset of patients with refractory symptoms requires surgical decompression of neural elements [2-4].

Various treatment modalities with variable success

rates have been used to manage LSS. To date, surgical decompression with laminectomy remains the gold standard treatment modality. However, choosing the optimal management strategy is usually guided by a set of clinical, radiological, and measurement indices. Currently, there is a major uncertainty and discrepancy in the methodology used, as well as inconsistent adoption of outcome measures across studies [5-7]. Various scales that assess symptom severity and functional disability have been established to evaluate treatment effectiveness as well as patients' pre- and post-treatment status. Such variations can impose huge limitations in predicting the efficacy and cost-effectiveness of different treatment paradigms. There-

Received Jun 17, 2017; Revised Jul 13, 2017; Accepted Jul 16, 2017

Corresponding author: Ankit I. Mehta

Department of Neurosurgery, University of Illinois at Chicago, 912 S. Wood St., M/C 799, Chicago, IL 60612, USA

Tel: +1-312-996-4842, Fax: +1-312-996-9018, E-mail: [ankitm@uic.edu](mailto:ankitm@uic.edu)

fore, having an optimal measurement index that correctly predicts treatment efficacy would significantly improve patient selection and outcomes.

Here, we review the various measurement indices used for outcome assessment in patients with LSS and delineate the major advantages and disadvantages of each index.

### Swiss Spinal Stenosis Questionnaire

The Swiss Spinal Stenosis (SSS) Questionnaire, also known as the “Swiss Spinal Stenosis measure,” the “Brigham Spinal Stenosis Questionnaire,” and the “Zurich Claudication Questionnaire” is specific to LSS and is used to supplement existing measures of disability and health status [8]. Studies have demonstrated its accuracy and reliability in measuring neurogenic claudication and walking capacity in patients with LSS [9,10]. The questionnaire was developed by Stucki et al. [8] in 1995 to address the symptoms and physical activity limitations within the prior month. It encompasses twelve questions related to two components: the first scale is for the assessment of symptom severity, whereas the second scale is for the assessment of functional disability caused by spinal stenosis [8].

The first scale (symptom severity) includes seven questions related to back and lower limb symptoms (pain, numbness, and weakness) and balance. The questions are divided into two domains: the pain (questions 1–4) and the neuroischemic domain (questions 5–7). Except for the question related to balance, each of the six questions are scored from 1 to 5, with 1 indicating the absence of symptom and 5 indicating very severe occurrences of symptom. The seventh question is related to balance and includes only three options scored as 1, 3, or 5 [8].

The second scale (functional disability) includes five questions (8–12) that primarily assess walking capacity. Each question is scored from 1 to 4, with higher scores indicating greater disability. Scores are calculated from each scale to provide an overall score, taking into consideration that no more than two responses should be missing. The extent of patient satisfaction after treatment can be measured using a third scale (six questions), which is not part of the primary questionnaire [8].

Various studies have shown that the SSS Questionnaire is a valid, reliable, and responsive tool for patients with LSS [8,11-14]. Unlike other scales, the SSS Scale is specific to LSS and has been identified as the “best and most specific outcome measure for spinal stenosis” by the North

American Spine Society guidelines for LSS [15].

Owing to its specificity, validity, and reliability, the SSS Questionnaire has been used in various studies and clinical trials. In addition, the scale focuses on neuroischemic symptoms, which are frequently encountered in patients with LSS. These types of questions are not present in most other scales. However, the scale has some limitations. Using Rasch analysis, the symptom severity scale fails to function as a unidimensional domain [16]. Making the scale unidimensional would require the deletion of questions three and six, thereby leading to a reduction in its reliability and power. Therefore, it is recommended to subdivide the first scale (symptom severity scale) into two separate subscales, which would address the general pain (questions 1–3) and neuroischemic symptoms (questions 4–7) [11,16]. Conversely, the physical disability scale performs as a unidimensional scale in Rasch analysis. However, question 11 is not clinically significant, and the physical disability scale would be appropriate for statistical analysis using the Rasch analysis model by excluding it [16].

### Quebec Back Pain Disability Questionnaire

The Quebec Back Pain Disability Questionnaire is a tool for measuring the pain and disability caused by back pain [17]. In the questionnaire, patients rate their difficulty with a set of tasks and activities. Patients rate each of the 20 activities on a scale from 0 (not difficult at all) to 5 (unable to perform). The overall score is calculated from the sum of individual scores from all 20 items, with a higher score indicating greater back pain and disability [17].

### Pain Disability Index

Pain Disability Index comprises seven questions used to measure the patient’s perceived level of disability across different daily activities, such as self-care, family responsibilities, and social activities. Scoring is done using a Visual Analog Scale (VAS), in which each of the seven questions is rated based on the severity of disability from 0 to 10 with a possible total score of 70. A higher score indicates more severe pain-related disability. The index’s construct validity and reliability have been proven using psychometric analysis [18].

## Oswestry Disability Index

Oswestry Disability Index (ODI) was developed by John O'Brien in 1976 is another tool to assess back pain-related disability [19]. It encompasses nine questions related to activities of daily living, such as walking and personal care, rated on Likert scales of 5 or 6 points. The score is calculated as a percentage of the total possible score of 53, with a greater score indicating greater back-related disability [19,20]. The ODI is one of the most commonly employed tools used by surgeons and is utilized in a large number of studies in the literature. Owing to its wide and common use, the ODI is a valid and reproducible instrument of back pain-related disability and has been used to evaluate walking in patients with LSS [21,22]. The ODI has been analyzed using the Rasch method [23,24].

There are four versions of the ODI, each having its own criteria [5,18,25-32]. Some versions contain questions related to sexual life and performance, which may not be acceptable in some cultures [33-36]. One of its major advantages is that it can be easily translated, and except for few questions, can be implemented in several cultures. The ODI has been translated into Danish, Dutch, French, German, Greek, Spanish, Swedish, and other languages [36-42].

The ODI is the preferred outcome metric for patients with severe symptoms. However, this preference may lead to a selection bias and incorrect reporting of outcomes [43]. Another issue related to the ODI is the time scale. ODI questions are developed to elicit a response about functioning "at the moment" the questions are asked. However, it is well known that symptom variation from one day to another is common among patients with LSS; thus, such specifications of time may overlook active issues the patient previously had but not currently experiencing when completing the questionnaire. This aspect affects the reliability of the test [12,43]. Additionally, the ODI scale does not consider the neuroischemic symptoms encountered by patients with LSS, thereby making its utility in patients experiencing such symptoms limited [43]. Conversely, the ODI scale has been shown to correlate most closely with the degree of patients' satisfaction after surgery for LSS [44].

## Short-Form Health Survey

The 36-item Short-Form Health Survey (SF-36) is an instrument developed to assess the overall health and

disability among patients with LSS [45]. It comprises 36 items, with each question categorized into one of the following scales: general health, physical functioning, social functioning, role limitation (physical limitation, role limitation), emotional, mental health, energy/fatigue, pain, and comparative health. A score from each scale is calculated independently using algorithms defined by the authors [45]. Generally, each scale is directly transformed into a 0–100 scale, with a lower score indicating more disability and a higher score indicating less disability. The SF-36 has been analyzed using the Rasch method [46-48]. SF-36 is widely used for the estimation of cost-effectiveness of treatments in health economics and constitutes a variable when calculating the quality-adjusted life year. Recently, a shorter version of the SF-36 was published with a reduction of the number of items from 36 to 12, for those interested in the physical and mental domains of the survey.

## Self-Paced Walking Test & Shuttle Walk Test

Self-Paced Walking Test (SPWT) assesses walking capacity by measuring the total distance the patient can continuously walk on a flat surface at a self-selected pace, until experiencing symptoms of LSS or reaching a limit of 30 minutes [49]. SPWT has demonstrated high retest reliability in patients with LSS [49]. As with other questionnaires that evaluate walking in patients with LSS, SPWT is mostly associated with measuring the capacity and not performance. The information gained from these questionnaires may provide an incomplete, and possibly misleading, estimate of a patient's level of disability. Therefore, adding daily activity monitors for a certain time may be advantageous, as it allows for measurement of both capacity and performance [50]. The objective nature of the SPWT is a major advantage because it eliminates any subjective bias like that observed with other outcome measures.

Shuttle Walk Test (SWT) is a ten-meter course performed on a flat surface without any obstacles. The patient listens to a tape recorder explaining the test prior to starting and is asked to walk ten meters within a specified time. The assessor counts the number of shuttles (laps) and maximum distance. The result of the test is given in meters (number of completed shuttles multiplied by 10) [51-53]. The advantage of this test is its simplicity, both in understandability as well as the minimal equipment involved. In one study, the authors reported that a change of 76 m would be required in SWT for 95% certainty that the

change was not a chance occurrence [12]. Therefore, using SWT to monitor the progression of symptoms among patients with LSS may be useful.

### Visual Analog Scale

VAS is a psychometric response scale for subjective characteristics or attitudes that cannot be directly measured [54,55]. It is a commonly used to evaluate pain in patients with LSS. The patients rate their level of pain by indicating a position along a continuous line between two end-points, wherein one end indicates “no pain” and the other indicates “worst possible pain.” VAS has been shown to be a very sensitive and reproducible test [56]. However, it remains a subjective measurement index and lacks specificity to LSS. Several studies have used VAS assessing pain related to various other etiologies [55,57]. Yet, VAS is still commonly used among surgeons dealing with LSS patients to estimate the severity of pain and intervene accordingly. A numerical system may be used with VAS, wherein 1 would indicate “no pain” and 10 would indicate “worst possible pain.”

### Japanese Orthopedic Association

Japanese Orthopedic Association (JOA) score was developed in 1976 to measure the functional status among patients with cervical myelopathy [58]. It was revised in 1994 to include assessment of shoulder and elbow functions [59]. The use of both scales has been reported in the literature, and they are considered the most frequently used scales in patients with cervical myelopathy. Despite its use in cervical myelopathy, the use of both scales has also been reported in a great number of studies in patients with LSS. The JOA scale comprises six domains, with each scored differently (Table 1) [58,60].

The scale was developed in Japan and was designed primarily for the Japanese population. Thus, one of the major limitations of this scale is that it evaluates motor dysfunction by assessing a patient’s ability to use chopsticks. The use of chopsticks is limited to East Asian cultures, including the Japanese, Korean, Chinese, and Vietnamese populations. Therefore, using this parameter to evaluate motor ability may not be applicable in the other parts of the world. This limitation of cultural differences in eating habits has already been addressed in previous reports [61,62].

**Table 1.** Japanese Orthopedic Association Scales and Scoring System<sup>a)</sup>

Scale	Score
Motor dysfunction in the upper extremities	0–4
Motor dysfunction in the lower extremities	0–4
Sensory function in the upper extremities	0–2
Sensory function in the trunk	0–2
Sensory function in the lower extremities	0–2
Bladder function	0–3

<sup>a)</sup>After scoring each domain, the sum of all domains is calculated for a maximum total score of 17 [58,60].

**Table 2.** Modified Japanese Orthopedic Association Scales and Scoring System<sup>a)</sup>

Scale	Score
Motor dysfunction in the upper extremities	0–7
Motor dysfunction in the lower extremities	0–5
Sensory function in the upper extremities	0–3
Bladder function	0–3

<sup>a)</sup>After scoring each domain, the sum of all domains is calculated for a maximum total score of 18. A mJOA score  $\geq 15$  indicates mild disease, 12–14 indicates moderate disease, and  $< 12$  indicates severe disease [63].

Like the ODI scale, the JOA score has been modified with three different versions reported in the literature, each with different scales and scoring systems [63–65]. Therefore, the potential of data misinterpretation and confusion across these different versions is not uncommon. The reliability of both JOA and modified JOA (mJOA) has been reported to be high [60,66]. However, the test-retest reliability of the mJOA has not yet been reported.

Out of the three modified versions, the most widely-used mJOA score is the one developed by Benzel et al. [63] in 1991. The scale comprises four domains, each scored differently (Table 2) [63]. One of the advantages of this scale is that it omits the question related to the use of chopsticks and replaces it with the use of spoons to evaluate the motor function of the upper limbs, thereby eliminating the issue associated with cultural differences [63].

This modification and others in the mJOA have not yet been validated; however, a recent study demonstrated that the domains and total scores of the JOA and the mJOA are strongly correlated with each other, even though the latter does not include scales for measuring sensory function in

**Table 3.** Oxford Claudication Score [71]

Question	Choices
1. Pain frequency: on an average, how often have you experienced pain in your back or buttock or pain that goes down your legs in the past month?	Not at all Less than once a week At least once a week Every day for at least a few minutes Every day for most of the day Every minute of the day
2. Total pain severity: on an average, how would you describe the worst pain you have had (pain in your back or buttock, or pain that goes down your legs) in the past month?	None Mild Moderate Severe Very severe Intolerable
3. Back pain severity: on an average, how would you describe the pain or discomfort in your back or buttocks in the past month?	None Mild Moderate Severe Very severe Intolerable
4. Leg pain severity: on an average, how would you describe the pain or discomfort in your legs or feet in the past month?	None Mild Moderate Severe Very severe Intolerable
5. Nerve symptom severity: on an average, how would you describe the numbness or tingling in your legs or feet in the past month?	None Mild Moderate Severe Very severe Intolerable
6. Leg weakness: on an average, how would you describe the strength in your legs, ankles, or feet in the past month?	None Mild Moderate Severe Very severe Intolerable
7. Balance: on an average, which statement best describes your steadiness when standing or walking in the past month?	I have had no problems with balance. I sometimes feel off-balance, but I am able to walk without any aid. I often feel off-balance, but I am able to walk with an aid. I am able to walk without an aid. I have difficulty walking despite using an aid. I cannot stand up.
8. Walking distance: on an average, in the past month, when you went for a walk, how far were you able to walk before your back or leg started giving trouble?	More than 2 miles or no limit More than 1/4 mile but less than 2 miles More than 100 yards but less than 1/4 mile More than 50 feet but less than 100 yards Less than 50 feet Not at all
9. Walking ability: on an average, which statement best describes your walking ability over the past month?	There is no limit to my walking ability. I can walk far enough to do everything I want to do. I am able to walk comfortably from my home to the shops or my transport. I am able to walk comfortably around the house. I am able to walk only from the bedroom to the bathroom or kitchen. I am not able to walk at all.

(Continued to the next page)

**Table 3.** Continued

Question	Choices
10. Walking speed: on an average, which statement best describes your walking over the past month?	I am able to walk at a normal speed. I walk slowly standing upright. I walk slowly bent forward. I have to stop and stand still when I walk. I have to stop and sit down when I walk. I cannot walk at all.

the trunk and lower extremities [67]. Yet, the Bland–Altman analysis indicated that these scoring systems are not interchangeable [67]. Finally, although the JOA and mJOA are widely reported in the LSS literature, both scales seem to be more specific for cervical myelopathy rather than LSS.

### Euro Quality of Life-5D

Euro Quality of Life-5D (EQ-5D) is composed of a descriptive scale and a VAS [68]. The former covers five domains: mobility, usual activities, self-care, anxiety/depression, and pain/discomfort. Each domain contains three options (no problems, some problems, severe problems). In the VAS, the patient is asked to rate his/her health on a vertical visual scale with options ranging from the “worst imaginable health state” to “best imaginable health state.” The EQ-5D scores for the US population range from –0.11 (worse than death) to 1.0 (full health), with a score of 0 indicating death [69].

One of the major advantages of this scale is that it measures both the economic as well as the clinical status of the patient. Thus, it provides the physician with a good estimate of the patient’s clinical condition and the cost-effectiveness of the treatment modality to be used [68,69]. Moreover, EQ-5D scores correlate well with the widely-used ODI scores [70]. Interestingly, the EQ-5D scales include a domain on “anxiety/depression,” which is not found in other scales and is very important in evaluating patients with LSS both pre- and postoperatively. However, as with other scales, the EQ-5D lacks specificity, as it can be used in several other spinal and non-spinal conditions and diseases. Owing to its lack of specificity, the conditions and environment of questionnaire administration may greatly influence and change the answers of the more generalized health questions, thereby affecting test reproducibility [70].

### Oxford Claudication Score

Oxford Claudication Score (OCS) comprises 10 questions covering three domains: pain (questions 1–4), ischemia (questions 5–7), and physical symptoms (questions 8–10) [71]. Each question is scored between 0 and 5, with higher scores correlating with worse disability (Table 3) [71]. The total score is expressed as a percentage of the maximum possible score. Like the SSS, the OCS includes questions related to symptoms experienced in the past month prior to presentation. The major advantage of the OCS is its specificity to LSS, particularly in terms of pain and walking difficulty [12,71].

### Conclusions

Each of the various outcome measures reported in the literature for the evaluation of patients with LSS have positive and negative attributes, thereby contributing to the variability in the observed success rates among various surgical and non-surgical interventions. Not a single outcome measure adequately captures all functional domains in patients with symptomatic LSS; therefore, a combination of questionnaires may be necessary to adequately capture the impact of LSS on a patient’s pain, disability, and quality of life. Thus, patients and physicians can potentially benefit from a single objective outcome measure that encompasses and addresses all these issues and can also monitor patients’ progression after treatment. Clinicians as well as third party payers can potentially benefit from a more rigid evaluation technique for LSS, thereby better ascertaining a judgment of functional improvement for the growing elderly population.



## Conflict of Interest

No potential conflict of interest relevant to this article was reported.

## References

1. Arbit E, Pannullo S. Lumbar stenosis: a clinical review. *Clin Orthop Relat Res* 2001;(384):137-43.
2. Ciol MA, Deyo RA, Howell E, Kreif S. An assessment of surgery for spinal stenosis: time trends, geographic variations, complications, and reoperations. *J Am Geriatr Soc* 1996;44:285-90.
3. Porter RW. Spinal stenosis and neurogenic claudication. *Spine (Phila Pa 1976)* 1996;21:2046-52.
4. Katz JN, Dalgas M, Stucki G, et al. Degenerative lumbar spinal stenosis: diagnostic value of the history and physical examination. *Arthritis Rheum* 1995;38:1236-41.
5. 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;23:2003-13.
6. Turner JA, Ersek M, Herron L, Deyo R. Surgery for lumbar spinal stenosis: attempted meta-analysis of the literature. *Spine (Phila Pa 1976)* 1992;17:1-8.
7. Siebert E, Pruss H, Klingebiel R, Failli V, Einhaupl KM, Schwab JM. Lumbar spinal stenosis: syndrome, diagnostics and treatment. *Nat Rev Neurol* 2009;5:392-403.
8. Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol* 1995;48:1369-78.
9. Fokter SK, Yerby SA. Patient-based outcomes for the operative treatment of degenerative lumbar spinal stenosis. *Eur Spine J* 2006;15:1661-9.
10. Zucherman JF, Hsu KY, Hartjen CA, et al. A multi-center, prospective, randomized trial evaluating the X STOP interspinous process decompression system for the treatment of neurogenic intermittent claudication: two-year follow-up results. *Spine (Phila Pa 1976)* 2005;30:1351-8.
11. Stucki G, Daltroy L, Liang MH, Lipson SJ, Fossel AH, Katz JN. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine (Phila Pa 1976)* 1996;21:796-803.
12. Pratt RK, Fairbank JC, Virr A. The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2002;27:84-91.
13. Tuli SK, Yerby SA, Katz JN. Methodological approaches to developing criteria for improvement in lumbar spinal stenosis surgery. *Spine (Phila Pa 1976)* 2006;31:1276-80.
14. Tomkins CC, Battie MC, Hu R. Construct validity of the physical function scale of the Swiss Spinal Stenosis Questionnaire for the measurement of walking capacity. *Spine (Phila Pa 1976)* 2007;32:1896-901.
15. North American Spine Society. Evidence-based clinical guidelines for multidisciplinary spine care: diagnosis and treatment of degenerative lumbar spinal stenosis. Burr Ridge (IL): North American Spine Society; 2007.
16. Comer CM, Conaghan PG, Tennant A. Internal construct validity of the Swiss Spinal Stenosis questionnaire: Rasch analysis of a disease-specific outcome measure for lumbar spinal stenosis. *Spine (Phila Pa 1976)* 2011;36:1969-76.
17. Kopec JA, Esdaile JM, Abrahamowicz M, et al. The Quebec Back Pain Disability Scale: measurement properties. *Spine (Phila Pa 1976)* 1995;20:341-52.
18. Tait RC, Chibnall JT, Krause S. The Pain Disability Index: psychometric properties. *Pain* 1990;40:171-82.
19. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980;66:271-3.
20. Katz JN, Lipson SJ, Chang LC, Levine SA, Fossel AH, Liang MH. Seven- to 10-year outcome of decompressive surgery for degenerative lumbar spinal stenosis. *Spine (Phila Pa 1976)* 1996;21:92-8.
21. Gunzburg R, Keller TS, Szpalski M, Vandeputte K, Spratt KF. Clinical and psychofunctional measures of conservative decompression surgery for lumbar spinal stenosis: a prospective cohort study. *Eur Spine J* 2003;12:197-204.
22. Herno A, Airaksinen O, Saari T, Pitkanen M, Manninen H, Suomalainen O. Computed tomography findings 4 years after surgical management of lumbar spinal stenosis: no correlation with clinical outcome. *Spine (Phila Pa 1976)* 1999;24:2234-9.
23. Page SJ, Shawaryn MA, Cernich AN, Linacre JM.

- Scaling of the revised Oswestry low back pain questionnaire. *Arch Phys Med Rehabil* 2002;83:1579-84.
24. Davidson M. Rasch analysis of three versions of the Oswestry Disability Questionnaire. *Man Ther* 2008;13:222-31.
  25. Hupli M, Sainio P, Hurri H, Alaranta H. Comparison of trunk strength measurements between two different isokinetic devices used at clinical settings. *J Spinal Disord* 1997;10:391-7.
  26. Boden SD. Outcome assessment after spinal fusion: why and how? *Orthop Clin North Am* 1998;29:717-28.
  27. Liebenson C. *Rehabilitation of the spine: a practitioner's manual*. Baltimore (MD): Williams & Wilkins; 1996.
  28. Baker D, Pynsent P, Fairbank J. The Oswestry Disability Index revisited. In: Roland M, Jenner JR, editors. *Back pain: new approaches to rehabilitation and education*. Manchester: Manchester University Press; 1989. p.174-86.
  29. Comparison of chiropractic and hospital outpatient management of low back pain: a feasibility study: report of a working group. *J Epidemiol Community Health* 1986;40:12-7.
  30. Meade TW, Dyer S, Browne W, Frank AO. Randomised comparison of chiropractic and hospital outpatient management for low back pain: results from extended follow up. *BMJ* 1995;311:349-51.
  31. Pynsent PB, Fairbank JC, Carr AJ. *Outcome measures in orthopaedics*. Oxford: Butterworth-Heinemann; 1993.
  32. Hudson-Cook N, Tomes-Nicholson K, Breen A. A revised Oswestry disability questionnaire. In: Roland M, Jenner JR, editors. *Back pain: new approaches to rehabilitation and education*. Manchester: Manchester University Press; 1989. p.187-204.
  33. Coleman RE, Houston S, Purohit OP, Rubens RD, Kandra A, Ford J. A randomised phase II study of oral pamidronate for the treatment of bone metastases from breast cancer. *Eur J Cancer* 1998;34:820-4.
  34. Poussa M, Schlenzka D, Seitsalo S, Ylikoski M, Hurri H, Osterman K. Surgical treatment of severe isthmic spondylolisthesis in adolescents. Reduction or fusion in situ. *Spine (Phila Pa 1976)* 1993;18:894-901.
  35. Purohit OP, Anthony C, Radstone CR, Owen J, Coleman RE. High-dose intravenous pamidronate for metastatic bone pain. *Br J Cancer* 1994;70:554-8.
  36. Tyrrell CT, Bruning PF, May-Levin F, et al. Pamidronate infusions as single-agent therapy for bone metastases: a phase II trial in patients with breast cancer. *Eur J Cancer* 1995;31A:1976-80.
  37. Manniche C, Asmussen K, Lauritsen B, Vinterberg H, Kreiner S, Jordan A. Low Back Pain Rating scale: validation of a tool for assessment of low back pain. *Pain* 1994;57:317-26.
  38. Drosy R, Marty M. Quality-of-life indexes for assessment of low-back-pain. *Rev Rhum* 1994;61:S44-8.
  39. Muller-Busch HC, Hoffmann P. Active music therapy for chronic pain: a prospective study. *Schmerz* 1997;11:91-100.
  40. Boscainos P, Sapkas G, Stilianesi E, et al. Clinical relevance of specific parameters isolated within the Oswestry and Roland-Morris functional disability scales. *J Bone Joint Surg [Br]* 1999;81-B:239.
  41. Kovacs FM, Abaira V, Lopez-Abente G, Pozo F. Neuro-reflexotherapy intervention in the treatment of non specified low back pain: a randomized, controlled, double-blind clinical trial. *Med Clin (Barc)* 1993;101:570-5.
  42. Skargren EI, Oberg BE, Carlsson PG, Gade M. Cost and effectiveness analysis of chiropractic and physiotherapy treatment for low back and neck pain: six-month follow-up. *Spine (Phila Pa 1976)* 1997;22:2167-77.
  43. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine (Phila Pa 1976)* 2000;25:2940-52.
  44. Herno A, Saari T, Suomalainen O, Airaksinen O. The degree of decompressive relief and its relation to clinical outcome in patients undergoing surgery for lumbar spinal stenosis. *Spine (Phila Pa 1976)* 1999;24:1010-4.
  45. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): I. conceptual framework and item selection. *Med Care* 1992;30:473-83.
  46. Haley SM, McHorney CA, Ware JE Jr. Evaluation of the MOS SF-36 physical functioning scale (PF-10): I. unidimensionality and reproducibility of the Rasch item scale. *J Clin Epidemiol* 1994;47:671-84.
  47. Stucki G, Daltroy L, Katz JN, Johannesson M, Liang MH. Interpretation of change scores in ordinal clinical scales and health status measures: the whole may not equal the sum of the parts. *J Clin Epidemiol* 1996;49:711-7.



48. Taylor WJ, McPherson KM. Using Rasch analysis to compare the psychometric properties of the Short Form 36 physical function score and the Health Assessment Questionnaire disability index in patients with psoriatic arthritis and rheumatoid arthritis. *Arthritis Rheum* 2007;57:723-9.
49. Tomkins CC, Battie MC, Rogers T, Jiang H, Petersen S. A criterion measure of walking capacity in lumbar spinal stenosis and its comparison with a treadmill protocol. *Spine (Phila Pa 1976)* 2009;34:2444-9.
50. Conway J, Tomkins CC, Haig AJ. Walking assessment in people with lumbar spinal stenosis: capacity, performance, and self-report measures. *Spine J* 2011;11:816-23.
51. Keell SD, Chambers JS, Francis DP, Edwards DF, Stables RH. Shuttle-walk test to assess chronic heart failure. *Lancet* 1998;352:705.
52. Payne GE, Skehan JD. Shuttle walking test: a new approach for evaluating patients with pacemakers. *Heart* 1996;75:414-8.
53. Singh SJ, Morgan MD, Scott S, Walters D, Hardman AE. Development of a shuttle walking test of disability in patients with chronic airways obstruction. *Thorax* 1992;47:1019-24.
54. Reips UD, Funke F. Interval-level measurement with visual analogue scales in Internet-based research: VAS Generator. *Behav Res Methods* 2008;40:699-704.
55. Adamchic I, Langguth B, Hauptmann C, Tass PA. Psychometric evaluation of visual analog scale for the assessment of chronic tinnitus. *Am J Audiol* 2012;21:215-25.
56. Grant S, Aitchison T, Henderson E, et al. A comparison of the reproducibility and the sensitivity to change of visual analogue scales, Borg scales, and Likert scales in normal subjects during submaximal exercise. *Chest* 1999;116:1208-17.
57. Bourdel N, Alves J, Pickering G, Ramilo I, Roman H, Canis M. Systematic review of endometriosis pain assessment: how to choose a scale? *Hum Reprod Update* 2015;21:136-52.
58. Japanese Orthopaedic Association. Japanese Orthopaedic Association scoring system for cervical spondylotic myelopathy. *Nippon Seikeigeka Gakkai Zasshi* 1976;50:18-9.
59. Japanese Orthopaedic Association. Scoring system for cervical myelopathy. *Nippon Seikeigeka Gakkai Zasshi* 1994;68:490-503.
60. 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;26:1890-4.
61. Chen FF. What happens if we compare chopsticks with forks?: the impact of making inappropriate comparisons in cross-cultural research. *J Pers Soc Psychol* 2008;95:1005-18.
62. Roh YH, Yang BK, Noh JH, Baek GH, Song CH, Gong HS. Cross-cultural adaptation and validation of the Korean version of the Michigan hand questionnaire. *J Hand Surg Am* 2011;36:1497-503.
63. Benzel EC, Lancon J, Kesterson L, Hadden T. Cervical laminectomy and dentate ligament section for cervical spondylotic myelopathy. *J Spinal Disord* 1991;4:286-95.
64. Chiles BW 3rd, Leonard MA, Choudhri HF, Cooper PR. Cervical spondylotic myelopathy: patterns of neurological deficit and recovery after anterior cervical decompression. *Neurosurgery* 1999;44:762-9.
65. Keller A, von Ammon K, Klaiber R, Waespe W. Spondylogenic cervical myelopathy: conservative and surgical therapy. *Schweiz Med Wochenschr* 1993;123:1682-91.
66. Bartels RH, Verbeek AL, Benzel EC, Fehlings MG, Guiot BH. Validation of a translated version of the modified Japanese orthopaedic association score to assess outcomes in cervical spondylotic myelopathy: an approach to globalize outcomes assessment tools. *Neurosurgery* 2010;66:1013-6.
67. Kato S, Oshima Y, Oka H, et al. Comparison of the Japanese Orthopaedic Association (JOA) score and modified JOA (mJOA) score for the assessment of cervical myelopathy: a multicenter observational study. *PLoS One* 2015;10:e0123022.
68. EuroQol Group. EuroQol: a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199-208.
69. Agency for Healthcare Research and Quality. Calculating the U.S. population-based EQ-5D index score. Rockville (MD): Agency for Healthcare Research and Quality; 2005 [cited 2017 May 10]. Available from: <http://www.ahrq.gov/rice/EQ5Dscore.htm>.
70. Mueller B, Carreon LY, Glassman SD. Comparison of the EuroQOL-5D with the Oswestry Disability Index, back and leg pain scores in patients with degenera-

- tive lumbar spine pathology. *Spine (Phila Pa 1976)* 2013;38:757-61.
71. Maken P, Fairbank J, Wanders L. Clinical assessment of lumbar spinal stenosis. *J Bone Joint Surg [Br]* 1998;30(2S):158.