Journal of Taibah University Medical Sciences (2016) 11(2), 91–97

Taibah University

Journal of Taibah University Medical Sciences

www.sciencedirect.com

Review Article

Why and how should we measure outcomes in spine surgery?

Alisson R. Teles, MD [a], Khalid I. Khoshhal, ABOS [b] and Asdrubal Falavigna, PhD [a,*]

[a] *Department of Neurosurgery, Universidade de Caxias do Sul, Laboratory of Clinical Studies and Basic Models of Spinal Disorders, Caxias do Sul, Brazil*
[b] *Department of Orthopedic Surgery, College of Medicine, Taibah University, Almadinah Almunawwarah, KSA*

Received 25 October 2015; revised 10 January 2016; accepted 10 January 2016; Available online 17 February 2016

الملخص

تهدف جراحات العمود الفقري للأمراض التنكسية إلى تخفيف الألم، والعجز، والتحسين النوعي لحياة المريض، والعودة للعمل. وقد أدى عدم التناسق بين نظرة الجراح من جهة لنتائج الفحوص الشعاعية من خلال وضع دمج الفقرات، وكفاية عمليات تخفيف الضغط على النخاع، ورضا المريض عن النتائج من جهة أخرى؛ إلى الحاجة لتقييم نواتج جراحات العمود الفقري من وجهة نظر المريض. وتعتبر تقارير المرضى للنواتج هي تقارير عن الوضع الصحي تؤخذ مباشرة من المرضى دون تدخل الأطباء. وعبر العقود الماضية طُور عدد من الأدوات العامة والخاصة لقياس تقارير المرضى للنواتج، كما تم التحقق من عدد منها على المصابين بآفات العمود الفقري. راجعت هذه الورقة أكثر أدوات قياس تقارير المرضى للنواتج استخداما لمرضى آفات العمود الفقري، مع التركيز على صفاتها وقابليتها للتطبيق ومعيار أقل اختلاف سريري يحدث فرقا.

الكلمات المفتاحية: العجز؛ نواتج؛ ألم؛ نوعية الحياة؛ العمود الفقري

Abstract

The objectives of spinal surgery for degenerative disorders are to reduce pain and disability and improve patients' quality of life while allowing an early return to work. The incongruence between surgeons' perspectives, findings of imaging exams in terms of fusion statuses or adequate decompression, and patients' satisfaction levels with treatment have underscored the need to evaluate outcomes of spinal surgery with a specific focus on patients' perspectives. Patient-reported outcomes (PRO) are reports on health status taken directly from patients without interference from physicians. In recent decades, several generic and disease-specific PRO instruments have been developed and validated in patients with spinal disorders. In this paper, we review the most commonly used PRO instruments in patients with spinal disorders, focussing on their characteristics, applicability and minimum clinically important differences.

Keywords: Disability; Outcomes; Pain; Quality of life; Spine

Introduction

Degenerative spinal disorders have a major impact on patients' quality of life. Surgical treatment is usually indicated for patients who do not respond to clinical therapy. Although the indications for surgery are usually based on the relationship between clinical findings, imaging exams and pain symptoms, several studies have demonstrated little correlation between these variables and the severity of the disease from patients' perspectives.[1–5]

The definition of a good surgical outcome depends on how success is assessed.[6] In the past, outcomes were commonly assessed based on surgeons' subjective views, and the results were ranked using terms such as "excellent", "good", "moderate" and "bad". The technical success of surgeries in terms of decompression and/or

* Corresponding address: Universidade de Caxias do Sul, Rua General Arcy da Rocha Nóbrega, 401/602, CEP: 95040-290, Caxias do Sul, RS, Brazil.
E-mail: asdrubalmd@gmail.com (A. Falavigna)
Peer review under responsibility of Taibah University.

ELSEVIER Production and hosting by Elsevier

 http://dx.doi.org/10.1016/j.jtumed.2016.01.003

fusion, as assessed by imaging studies during patients' follow-up visits, also influenced these classification schemes. However, surgeons' perspectives and results of imaging exams frequently do not correlate with patient satisfaction.[1–3,7]

In patients with degenerative spinal disorders, the main objective of treatment is to improve health-related quality of life by reducing pain and disability. In this sense, the best measurement of treatment quality should be the patient's opinion of the results using patient-reported outcomes (PRO) instruments. In recent decades, several questionnaires have been developed and validated for the clinical and functional evaluation of spinal treatment outcomes. These instruments, in addition to assessing outcomes from patients' perspectives, can be used to select patients for whom surgery is appropriate, as preoperative values often correlate with the success of treatment.[8–11]

In this article, the authors provide a thorough overview of the main PRO instruments commonly used in spinal disorders. Special attention is given to their clinical applications and the interpretation of their results. Finally, we report our experiences since 2006 in using PRO in routine spine care.

Assessing patient-reported outcomes in spinal disorders

PRO are health status reports taken directly from patients without interpretation by clinicians.[12] In recent years, many generic and disease-specific PRO instruments have been validated worldwide to assess the outcomes of spine-related treatments.[10,11,13–16] The evaluation of PRO in spine care provides useful information for quality improvement and comparisons of effectiveness.[17] Table 1 summarizes the main applications of the data collected from PRO instruments.

PRO instruments provide results as numerical scores. Pre- and post-operative scores can be compared to detect changes. Validation studies comparing changes in numerical scores with standard categorical scales (i.e., anchor-based validation) have identified the minimal changes in the instruments that represent clinical significance for patients.[18,19] This concept is called the minimum clinical important difference (MCID),[19] and it is generally used as a parameter to establish the clinical effectiveness of a treatment. In other words, the MCID is the smallest difference in the instrument score that is significant for patients. It is important to understand that these values depend on both the instrument type and the pathology, as well as the patient's history of previous spinal surgeries.[18,20,21]

Table 1: Examples of patient-reported outcomes instrument use in spine care.

Evaluating the burden of disease.
Monitoring treatment outcomes.
Facilitating patient-physician communication.
Allowing results of treatments to be compared between different centres.
Measuring treatment quality for health economic analyses in spine care.
Can be used by health care provider organizations for benchmarking and quality improvement.

Another important factor is the difference between a statistically significant and a clinically significant difference in the PRO instruments: a statistically significant difference is not always clinically relevant from a patient's perspective.[22,23]

Given the great variety of assessment tools available in the literature, it is essential for spine surgeons to have the basic knowledge needed to choose and evaluate the results of the most appropriate tools for each disease or outcome of interest. Generally, the quality of instruments is assessed through their reliability, validity, responsiveness, acceptability, feasibility and ceiling and floor effects.[19] Reliability is the instrument's consistency in producing similar results in similar situations; variation is the error measure. The smaller the variation, the greater the reliability of the instrument. Validity refers to the degree to which an instrument actually measures the variable of interest. Responsiveness is the sensitivity of a test in detecting clinically relevant changes in a given scenario, notwithstanding its size. Acceptability relates to how appropriate the questions in the instrument are for the patients answering them; this refers to characteristics related to the time needed to complete the questionnaire, the way the instrument is administered, the proportion of missing data, and the difficulty in understanding and interpreting the questions. Feasibility refers to the ease of administering the instrument and processing its results; for example, self-administered and short instruments are more practical than semi-structured interviews and extensive questionnaires. Finally, ceiling and floor effects refer to the ability of a test to discriminate groups of patients and to detect changes in the disease that is being evaluated; for example, the ability to distinguish between patients with moderate versus severe disabilities arising from a particular disease.

Below, we review the PRO instruments most commonly used for patients with spinal disorders (Table 2). Special attention is given to their clinical applications and the interpretation of their results.

Pain

The assessment of pain intensity must be conducted using standardized scales due to the subjectivity of the symptom.[24,25] Visual and numerical scales have been used to assess pain intensity. Some examples of visual analogue scales are[1]: visual analogue scales of cups, which relate the amount of liquid inside a glass to the intensity of the pain perceived by the patient[2]; visual analogue scales of faces, in which the patient identifies the face that best represents his/her perception of the pain at any given time[3]; metric visual analogue scales, in which the patient marks a point on a 10-cm horizontal line that represents no pain on one end and the maximum possible pain on the other. The evaluator then measures the distance from the beginning of the line to the marked point and computes the pain intensity in millimetres.

The numerical rating scale of pain intensity is the most widely used pain scale for spinal surgeries.[10,11,26] The patient

Table 2: Commonly used generic and specific patient-reported outcomes instruments.

Instrument	Questions	Score range	MCID	Observations
SF36	36	0–100	4.9[18]	The most-used generic instrument for the assessment of health-related quality of life. The SF6D can be derived from the SF36, and it is widely used as utility measure. 2 components: physical and mental.
EQ5D	5	−0.594 – 1	–	A utility measure. Evaluates five areas: mobility, self-care, usual activities (work, study, housework, family or leisure activities), pain/discomfort and anxiety/depression.
ODI	10	0–100	10–12.8[18,30]	Disability in lumbar disease. Scores: 0–20, minimal disability; 21–40, moderate disability; 41–60, severe disability; 61–80, crippling back pain; 81–100, patients are either bed-bound or are exaggerating their symptoms.
NDI	10	0–100	8[37]	Disability in cervical disease. Scores: 0–10, no disability; 11–30, minimal disability; 31–50, moderate disability; 51–70, severe disability; above 71, total disability.
NRS of pain	1	0–10	1 to 2 points[18]	Applied separately for axial or radicular pain.
BDI	21	0–63	–	Depression. Scores: 0–13, no depression or minimal symptoms; 14–19, mild depression; 20–28, moderate depression; 29–63, severe depression.
HADS	14	0–21	–	Depression and anxiety. Composed of two subscales of depression and anxiety, each with 7 questions. The depression subscale has a high correlation with BDI scores.

MCID: minimal clinically important difference; SF36: short-form health survey 36 questions; SF6D: short-form health survey 6D; EQ5D: Euro Quality of Life 5D; ODI: Oswestry Disability Index; NDI: Neck Disability Index; NRS: Numerical Rating Scale; BDI: Beck Depression Inventory; HADS: Hospital Anxiety and Depression Scale.

lists the intensity of their pain on a scale from zero (no pain) to 10 (maximum possible pain). The scale should be applied separately for axial pain (cervical pain and lower back pain) and radicular pain (leg or arm pain). The MCIDs for the metric and numerical visual analogue scales of pain are 15–20 mm and 1–2 points, respectively.[18]

Disability

The assessment of disability is performed using specific instruments according to the topography of the pathology being evaluated (i.e., cervical versus lower back) and the type of disease or symptoms (i.e., pain versus myelopathy).

The Oswestry Disability Index (ODI) is the scale most commonly used to assess disability in lumbar spine pathologies.[27,28] The scale consists of 10 questions with six possible answers whose values range from 0 to 5, with the total score calculated and presented on a scale from zero to 100: 0–20 indicates minimal disability; 21–40 indicates moderate disability; 41–60 indicates severe disability; 61–80 indicates crippling back pain; and 81–100 indicates that the patient is either bed-bound or exaggerating their symptoms.[28] The scale assesses the domains of pain, personal care (dressing and bathing), lifting, walking, sitting, standing, sleeping, sexuality, socializing and the ability to travel. The clinically relevant minimum difference is 10–12.8 points in the total score.[18,29]

Other instruments that assess disability due to lower back disorders are the Roland–Morris Disability Questionnaire (RMDQ)[30] and the Quebec Back Pain Disability Scale (QBPDS).[31] The RMDQ is more appropriate than the ODI for the assessment of patients with milder disabilities.[32] It consists of 24 statements that assess the impact of lower back pain in daily life and in work activities. The results vary from zero (no impairment) to 24 (severe impairment), and the MCID varies from 3 to 6 points, depending on the author.[32] The QBPDS consists of 20 items describing the difficulty of performing physical activities of mild intensity. The possible scores for each item range from zero (no difficulty) to 5 (maximum inability to perform activities), and the final score can range from zero to 100. The MCID is 20 points in the total score.[20]

In cervical spine diseases, the most widely used scale is the Neck Disability Index,[11,33–35] which is composed of 10 questions with six possible answers to each ranging from 0 to 5. The total score is calculated and presented on a scale from zero (no disability) to 100 (maximum disability). Scores from zero to 10 represent no disability; 11 to 30, minimal disability; 31 to 50, moderate disability; 51 to 70, severe disability; and above 71, total disability.[35] A difference of 8 points is considered clinically relevant.[36]

Another scale used in several studies is the Neck Pain and Disability Scale (NPDS), which consists of 20 items scored on a visual analogue scale from zero (normal function) to 5 (maximum disability due to neck pain).[33] The instrument assesses functional disabilities in four areas: general neck problems, pain intensity, emotional effects or changes in mood resulting from neck pain, and effects on daily activities.

Outcome assessments for cervical myelopathy are usually performed with scales that measure the severity of symptoms. The most widely used are the Nurick Grade[37] and the modified Japanese Orthopedic Association scale (mJOA).[37,38] The Nurick Grade classifies patients on a scale from zero to five, where zero refers to patients without evidence of myelopathy, 1 indicates patients with clinical signs of myelopathy but without gait abnormalities, 2 indicates patients with some gait impairments but

without impacts on daily activities, 3 indicates patients with severe gait impairments but who are still able to walk without assistance, 4 indicates patients who need assistance to walk and 5 indicates patients who use wheelchairs.[37] The mJOA assesses the motor and sensory functions of the upper and lower limbs, as well as trunk and bladder function. The final score ranges from zero (maximum disability) to 17 (normal function). Scores lower than 12 are considered indicative of severe disability.[37]

Quality of life

The most widely used instrument for assessing health-related quality of life is the Short Form 36 Health Survey Questionnaire (SF36).[39] It is a multidimensional generic instrument composed of 36 items, and it is validated for several diseases, including diseases of the spine.[10,11] The SF36 has two components, physical and mental, and eight domains. The physical component evaluates the domains of functional capacity, physical aspects, pain and general health. The mental component assesses vitality, social functioning, emotional aspects and mental health. The results are transformed into a scale from zero to 100, with 50 points as the average measured in the general population, a standard deviation of 10 points, and an MCID of 4.9 points in the component scores.[18] Shorter versions of the instrument have been used in several studies (SF12 and SF6) with good reliability and reproducibility compared with the SF36.[39–41]

Other examples of generic instruments are the World Health Organization Quality of Life Assessment – short version[42] and the Euro Quality of Life – 5 Dimensions (EQ5D).[43] The EQ5D is a generic instrument that is widely used to measure preference-based health status for health economic analyses. This short instrument evaluates five areas: mobility, self-care, usual activities (work, study, housework, and family or leisure activities), pain/discomfort and anxiety/depression.

Psychological aspects

The evaluation of psychological aspects is critical in the preoperative period and during follow-up care for patients with spinal disorders.[15] The prevalence of mood disorders is very high in these patients.[11] Moreover, it has been clearly established that there is a relationship between the presence of depression and anxiety in the preoperative period and higher chances of failure after surgical treatment for degenerative diseases of the spine.[44,45] Additionally, we recently demonstrated that depression can be identified in 65% of patients following spinal surgery and that improvement in depression is associated with positive clinical outcomes.[46]

The most widely used instrument for the assessment of depressive symptoms is the Beck Depression Inventory (BDI), which consists of 21 questions with scores ranging from zero to 3.[47] The instrument assesses the following symptoms: sadness, pessimism, sense of failure, feelings of guilt, self-dislike, suicidal thoughts, weight loss, work inhibition, sleep disturbances, fatigability, and loss of libido. Scores range from zero to 63. Scores of 0–13 indicate no depression or minimal symptoms, scores of 14–19 indicate mild depression, scores of 20–28 indicate moderate depression, and scores of 29–63 indicate severe depression.[11]

The Zung Self-Rating Depression Scale is also commonly used for patients with spinal disorders.[48–50] The questionnaire consists of 20 statements related to specific characteristics of depression. These are evaluated using Likert scale responses. The final score is the sum of the scores for each question. The minimum possible score is 20 and the maximum possible score is 80: scores of 20–49 indicate no depression, scores of 50–59 indicate mild depression, scores of 60–69 indicate moderate depression, and scores of 70–80 indicate severe depression.

The Hospital Anxiety and Depression Scale (HADS) consists of 14 multiple-choice questions and two subscales. Each subscale for depression and anxiety consists of 7 questions, and total scores can range from zero to 21. Using 8–9 as the cut-off points for both subscales, the sensitivities and specificities are 93.7% and 72.6% for anxiety and 84.6% and 90.3% for depression, respectively.[51] The depression subscale has a high correlation with the BDI.[15]

Another important factor when evaluating the results of spinal disease treatment is the assessment of fears and beliefs related to physical activity and work. One of the most commonly used instruments for measuring these responses is the Fear-Avoidance Beliefs Questionnaire, which contains 16 questions related to physical activity (FABQ-FA, 5 questions) and work (FABQ-W, 11 questions).[52] This questionnaire uses Likert scale responses with seven choices ranging from zero (completely disagree) to 6 (completely agree). Validation studies have shown that questions 13, 14 and 16 are redundant and that issues 1 and 8 have a very low correlation with the total scale; therefore, they are excluded from the final score.[53] The FABQ-FA score can range from 0 to 24, and the FABQ-W score can range from 0 to 42 points. The higher the score, the greater the fear-avoidance beliefs related to work and physical activity.[53, 54]

Use of PRO in routine spine care: the experience of the *Universidade de Caxias do Sul*'s Spine Surgery Group

In 2006, we started our prospective clinical registry with patients selected for surgery by the *Universidade de Caxias do Sul*'s Spine Surgery Group. All patients undergoing elective spinal surgery for degenerative diseases are invited to enrol in this prospective clinical registry. Preoperative data and follow-up data (collected at 30 days, 6 months, and annually) are collected, including demographic information, PRO, radiological results, and costs. The full protocol has been reported in detail elsewhere.[11] The PRO instruments examine the following domains: pain (numerical rating scales of axial and radicular pain), spine-related disability (using the ODI for lumbar and the NDI for cervical), quality of life (SF36), and psychological aspects (BDI, HADS, and FABQ). For utility measures, we can extract SF6D values from the SF36 using specific Brazilian metrics.[41] All of the data are collected and managed using a database specifically built for this registry.

Our initial experience with the clinical registry and routine collection of PRO is encouraging. From 2006 to early 2015,

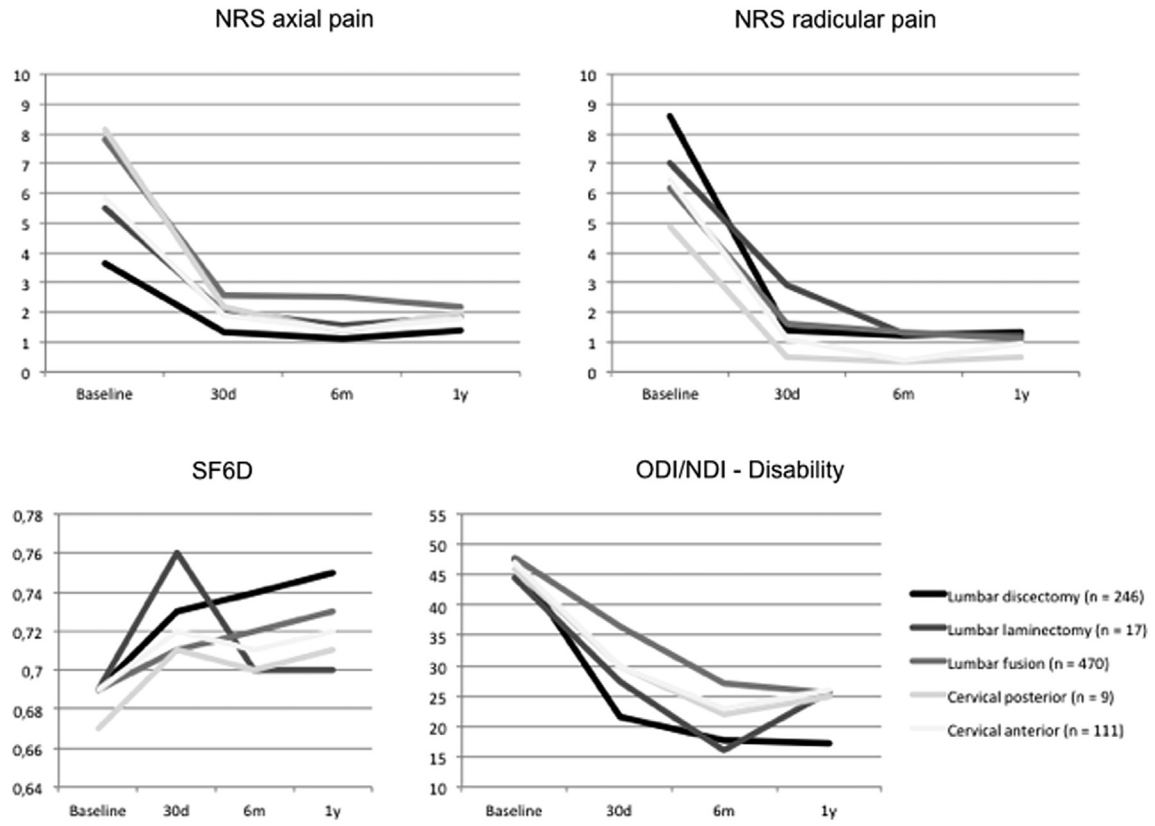

Figure 1: Improvements in patient-reported outcomes after commonly performed spinal procedures by the Universidade de Caxias do Sul Spine Surgery Group – Brazil. Note: ODI: Oswestry Disability Index, NDI: Neck Disability Index, NRS: Numerical Rating Scale, SF6D: Short-Form 6 Dimensions.

843 patients completed 1-year of follow-up care. As previously reported,[11] patients' perceptions of completing PRO instruments have been very positive. Using a specific database, we can assess clinical outcomes, pathologies and surgeries from patient to patient (Figure 1). This strengthens the patient–physician relationship, facilitates communication between the two parties and enables comparisons with different centres' reports in the literature.

As the results of PRO instruments are also used for research purposes, all of the patients invited to participate must sign an informed consent form. Before starting this registry, the protocol was submitted and approved by the local ethics committee. To date, several research articles have been published using this database.[15,46,55–57] We understand that the opportunity to share routine experiences is important to increasing research in developing countries.[58,59]

The value of a health intervention is based on its quality divided by its cost. In patients with degenerative spinal disorders, the main objective of treatment is to improve HRQOL by reducing pain and disability. In this sense, the best measurements of the quality of treatments should be patients' own reports regarding their outcomes, that is, patient-reported outcomes (PRO). The majority of the studies on the comparative effectiveness of spine care have been conducted in developed countries. However, due to the high prevalence and social and economic burdens of spinal disorders all over the world, comparative effectiveness research should also be conducted in developing countries, where health care resources are often distributed without adequate rationale and many health services are often underfunded.[57] Thus, it is essential for spine surgeons, as health care providers, to begin collecting PRO. It is expected that, as in developed countries, there will be a shift in paradigm towards value-based health care in the near future.[60]

Our group recently published the first health economic analysis of spinal surgery in Brazil.[56] Using data from PRO collected routinely in patients undergoing surgery for lumbar disc herniation, we demonstrated that lumbar discectomy is a very cost-effective treatment for patients with lumbar disc herniation refractory to clinical treatment from private and public health care perspectives.[56] Thus, in addition to serving to evaluate patients' outcomes from their own perspectives, the routine collection of PRO instruments for patients with spinal disorders can be used for health economic analyses. We strongly believe that these initiatives are extremely important in the contemporary paradigm of value-based health care, principally in developing countries, where routine assessments of PRO are performed by a very limited proportion of spine care groups.

Final considerations

Spinal disorder treatment outcomes should be assessed from patients' perspectives. The objectives of treatments are to reduce pain and disability and to increase patients' quality

of life and occupational capacity. There are many validated PRO instruments available that measure different clinical and functional outcomes. It is essential for spine care providers to be aware of and routinely use PRO instruments as needed. Outcome assessments made using PRO instruments help spine surgeons to better manage their patients and monitor treatment progress.

Authors' contribution

AF conceived and designed the study, conducted research, provided research materials, and collected and organized data. KIK analyzed, interpreted data, and revised it critically. ART wrote initial and final draft of article, and provided all the revision. All authors have critically reviewed and approved the final draft and are responsible for the content and similarity index of the manuscript. The paper submitted is an original article and there is no plagiarism.

Conflicts of interest

The authors have no conflict of interest to declare.

Acknowledgements

The authors would like to acknowledge the support of AOSpine LA and AOTrauma ME.

References

1. Righesso O, Falavigna A, Avanzi O. Correlation between persistent neurological impairment and clinical outcome after microdiscectomy for treatment of lumbar disc herniation. **Neurosurgery 2012**; 70(2): 390–396. discussion 6–7.
2. Djurasovic M, Carreon LY, Crawford 3rd CH, Zook JD, Bratcher KR, Glassman SD. The influence of preoperative MRI findings on lumbar fusion clinical outcomes. **Eur Spine J 2012**; 21(8): 1616–1623.
3. Djurasovic M, Glassman SD, Dimar 2nd JR, Howard JM, Bratcher KR, Carreon LY. Does fusion status correlate with patient outcomes in lumbar spinal fusion? **Spine (Phila Pa 1976) 2011**; 36(5): 404–409.
4. Goni VG, Hampannavar A, Gopinathan NR, Singh P, Sudesh P, Logithasan RK, et al. Comparison of the oswestry disability index and magnetic resonance imaging findings in lumbar canal stenosis: an observational study. **Asian Spine J 2014**; 8(1): 44–50.
5. Weber C, Giannadakis C, Rao V, Jakola AS, Nerland U, Nygaard OP, et al. Is there an association between radiological severity of lumbar spinal stenosis and disability, pain, or surgical outcome?: a multicenter observational study. **Spine (Phila Pa 1976) 2016 Jan**; 41(2): E78–83.
6. Chapman JR, Norvell DC, Hermsmeyer JT, Bransford RJ, DeVine J, McGirt MJ, et al. Evaluating common outcomes for measuring treatment success for chronic low back pain. **Spine (Phila Pa 1976) 2011**; 36(21 Suppl): S54–S68.
7. Haefeli M, Elfering A, Aebi M, Freeman BJ, Fritzell P, Guimaraes Consciencia J, et al. What comprises a good outcome in spinal surgery? A preliminary survey among spine surgeons of the SSE and European spine patients. **Eur Spine J 2008**; 17(1): 104–116.
8. Carreon LY, Glassman SD, Djurasovic M, Dimar JR, Johnson JR, Puno RM, et al. Are preoperative health-related quality of life scores predictive of clinical outcomes after lumbar fusion? **Spine (Phila Pa 1976) 2009**; 34(7): 725–730.
9. Falavigna A, Righesso O, Teles AR. Re: are preoperative health-related quality of life scores predictive of clinical outcomes after lumbar fusion? **Spine (Phila Pa 1976) 2009**; 34(23): 2596–2597.
10. Falavigna A, Teles AR, Braga GLd, Barazzetti DO, Lazzaretti L, Tregnago AC. Instrumentos de avaliação clínica e funcional em cirurgia da coluna vertebral. **Coluna/Columna 2011**; 10: 62–67.
11. Falavigna A, Righesso O, Teles AR. Avaliação clínica e funcional no pré-operatório de doenças degenerativas da coluna vertebral. **Coluna/Columna 2009**; 8(3): 245–253.
12. Greenhalgh J. The applications of PROs in clinical practice: what are they, do they work, and why? **Qual Life Res 2009**; 18(1): 115–123.
13. Mannion AF, Elfering A. Predictors of surgical outcome and their assessment. **Eur Spine J 2006**; 15(Suppl. 1): S93–S108.
14. Mannion AF, Elfering A, Staerkle R, Junge A, Grob D, Dvorak J, et al. Predictors of multidimensional outcome after spinal surgery. **Eur Spine J 2007**; 16(6): 777–786.
15. Falavigna A, Righesso O, Teles AR, Baseggio N, Velho MC, Ruschel LG, et al. Depression subscale of the hospital anxiety and depression scale applied preoperatively in spinal surgery. **Arq Neuropsiquiatr 2012**; 70(5): 352–356.
16. 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.
17. Azad TD, Kalani M, Wolf T, Kearney A, Lee Y, Flannery L, et al. Building an electronic health record integrated quality of life outcomes registry for spine surgery. **J Neurosurg Spine 2015**: 1–10.
18. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the oswestry disability index, medical outcomes study questionnaire short form 36, and pain scales. **Spine J 2008**; 8(6): 968–974.
19. Copay AG, Subach BR, Glassman SD, Polly Jr DW, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. **Spine J 2007**; 7(5): 541–546.
20. Ostelo RW, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. **Spine (Phila Pa 1976) 2008**; 33(1): 90–94.
21. Parker SL, Godil SS, Shau DN, Mendenhall SK, McGirt MJ. Assessment of the minimum clinically important difference in pain, disability, and quality of life after anterior cervical discectomy and fusion: clinical article. **J Neurosurg Spine 2013**; 18(2): 154–160.
22. Parker SL, Mendenhall SK, Shau DN, Adogwa O, Anderson WN, Devin CJ, et al. Minimum clinically important difference in pain, disability, and quality of life after neural decompression and fusion for same-level recurrent lumbar stenosis: understanding clinical versus statistical significance. **J Neurosurg Spine 2012**; 16(5): 471–478.
23. Mattei TA. "Statistically significant" does not necessarily mean 'clinically different' on pain/quality of life scales: opportune remarks on clinical outcomes measures in cervical spondylotic myelopathy. **Neurosurgery 2012**; 71(2): E518–E521. author reply E21–2.
24. Wang P, Zhang J, Liao W, Zhao L, Guo Y, Qiu Z, et al. Content comparison of questionnaires and scales used in low back pain based on the international classification of functioning, disability and health: a systematic review. **Disabil Rehabil 2012**; 34(14): 1167–1177.
25. Longo UG, Loppini M, Denaro L, Maffulli N, Denaro V. Rating scales for low back pain. **Br Med Bull 2010**; 94: 81–144.

26. Scheer JK, Smith JS, Clark AJ, Lafage V, Kim HJ, Rolston JD, et al. Comprehensive study of back and leg pain improvements after adult spinal deformity surgery: analysis of 421 patients with 2-year follow-up and of the impact of the surgery on treatment satisfaction. **J Neurosurg Spine 2015**; 22(5): 540–553.
27. Fairbank JC. Why are there different versions of the oswestry disability index? **J Neurosurg Spine 2014**; 20(1): 83–86.
28. Fairbank JC, Pynsent PB. The oswestry disability index. **Spine (Phila Pa 1976) 2000**; 25(22): 2940–2952. discussion 52.
29. Hagg O, Fritzell P, Nordwall A, Swedish Lumbar Spine Study G. The clinical importance of changes in outcome scores after treatment for chronic low back pain. **Eur Spine J 2003**; 12(1): 12–20.
30. Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. **Spine (Phila Pa 1976) 1983**; 8(2): 141–144.
31. Rodrigues MF, Michel-Crosato E, Cardoso JR, Traebert J. Psychometric properties and cross-cultural adaptation of the Brazilian Quebec back pain disability scale questionnaire. **Spine (Phila Pa 1976) 2009**; 34(13): E459–E464.
32. Roland M, Fairbank J. The Roland-Morris disability questionnaire and the oswestry disability questionnaire. **Spine (Phila Pa 1976) 2000**; 25(24): 3115–3124.
33. Cook C, Richardson JK, Braga L, Menezes A, Soler X, Kume P, et al. Cross-cultural adaptation and validation of the Brazilian Portuguese version of the neck disability index and neck pain and disability scale. **Spine (Phila Pa 1976) 2006**; 31(14): 1621–1627.
34. Chotai S, Parker SL, Sivaganesan A, Godil SS, McGirt MJ, Devin CJ. Quality of life and general health after elective surgery for cervical spine pathologies: determining a valid and responsive metric of health state utility. **Neurosurgery 2015**; 77(4): 553–560.
35. Sterling M, Rebbeck T. The neck disability index (NDI). **Aust J Physiother 2005**; 51(4): 271.
36. Carreon LY, Glassman SD, Campbell MJ, Anderson PA. Neck disability index, short form-36 physical component summary, and pain scales for neck and arm pain: the minimum clinically important difference and substantial clinical benefit after cervical spine fusion. **Spine J 2010**; 10(6): 469–474.
37. Kalsi-Ryan S, Singh A, Massicotte EM, Arnold PM, Brodke DS, Norvell DC, et al. Ancillary outcome measures for assessment of individuals with cervical spondylotic myelopathy. **Spine (Phila Pa 1976) 2013**; 38(22 Suppl 1): S111–S122.
38. Wilson JR, Fehlings MG, Kalsi-Ryan S, Shamji MF, Tetreault LA, Rhee JM, et al. Diagnosis, heritability, and outcome assessment in cervical myelopathy: a consensus statement. **Spine (Phila Pa 1976) 2013**; 38(22 Suppl 1): S76–S77.
39. Ware Jr JE. SF-36 health survey update. **Spine (Phila Pa 1976) 2000**; 25(24): 3130–3139.
40. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. **J Health Econ 2002**; 21(2): 271–292.
41. Cruz LN, Camey SA, Hoffmann JF, Rowen D, Brazier JE, Fleck MP, et al. Estimating the SF-6D value set for a population-based sample of Brazilians. **Value Health 2011**; 14(5 Suppl 1): S108–S114.
42. The World Health Organization Quality of Life Assessment (WHOQOL): development and general psychometric properties. **Soc Sci Med 1998**; 46(12): 1569–1585.
43. Coretti S, Ruggeri M, McNamee P. The minimum clinically important difference for EQ-5D index: a critical review. **Expert Rev Pharmacoecon Outcomes Res 2014**; 14(2): 221–233.
44. Chaichana KL, Mukherjee D, Adogwa O, Cheng JS, McGirt MJ. Correlation of preoperative depression and somatic perception scales with postoperative disability and quality of life after lumbar discectomy. **J Neurosurg Spine 2011**; 14(2): 261–267.
45. Pakarinen M, Vanhanen S, Sinikallio S, Aalto T, Lehto SM, Airaksinen O, et al. Depressive burden is associated with a poorer surgical outcome among lumbar spinal stenosis patients: a 5-year follow-up study. **Spine J 2014**; 14(10): 2392–2396.
46. Falavigna A, Righesso O, Teles AR, Conzati LP, Bossardi JB, da Silva PG, et al. Responsiveness of depression and its influence on surgical outcomes of lumbar degenerative diseases. **Eur J Orthop Surg Traumatol 2015**; 25(Suppl 1): S35–S41.
47. Choi Y, Mayer TG, Williams MJ, Gatchel RJ. What is the best screening test for depression in chronic spinal pain patients? **Spine J 2014**; 14(7): 1175–1182.
48. Adogwa O, Parker SL, Shau DN, Mendenhall SK, Aaronson OS, Cheng JS, et al. Preoperative Zung depression scale predicts outcome after revision lumbar surgery for adjacent segment disease, recurrent stenosis, and pseudarthrosis. **Spine J 2012**; 12(3): 179–185.
49. Adogwa O, Verla T, Thompson P, Penumaka A, Kudyba K, Johnson K, et al. Affective disorders influence clinical outcomes after revision lumbar surgery in elderly patients with symptomatic adjacent-segment disease, recurrent stenosis, or pseudarthrosis: clinical article. **J Neurosurg Spine 2014**; 21(2): 153–159.
50. Adogwa O, Parker SL, Shau DN, Mendenhall SK, Bydon A, Cheng JS, et al. Preoperative Zung depression scale predicts patient satisfaction independent of the extent of improvement after revision lumbar surgery. **Spine J 2013**; 13(5): 501–506.
51. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the hospital anxiety and depression scale. An updated literature review. **J Psychosom Res 2002**; 52(2): 69–77.
52. Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. **Pain 1993**; 52(2): 157–168.
53. Abreu AM, Faria CD, Cardoso SM, Teixeira-Salmela LF. [The Brazilian version of the fear avoidance beliefs questionnaire]. **Cad Saude Publica 2008**; 24(3): 615–623.
54. Staerkle R, Mannion AF, Elfering A, Junge A, Semmer NK, Jacobshagen N, et al. Longitudinal validation of the fear-avoidance beliefs questionnaire (FABQ) in a Swiss-German sample of low back pain patients. **Eur Spine J 2004**; 13(4): 332–340.
55. Falavigna A, Righesso O, Traynelis VC, Teles AR, da Silva PG. Effect of deep wound infection following lumbar arthrodesis for degenerative disc disease on long-term outcome: a prospective study: clinical article. **J Neurosurg Spine 2011**; 15(4): 399–403.
56. Falavigna A, Scheverin N, Righesso O, Teles AR, Gullo MC, Cheng JS, et al. Economic value of treating lumbar disc herniation in Brazil. **J Neurosurg Spine 2015**: 1–7.
57. Teles AR, Righesso O, Gullo MC, Ghogawala Z, Falavigna A. Perspective of value-based management of spinal disorders in Brazil. **World Neurosurg 2016 March**; 87: 346–354.
58. Falavigna A. Strategy to increase research in Latin America: project on education in research by AOSpine Latin America. **Glob Spine J 2015**; 5(4): 264–265.
59. Falavigna A, Botelho RV, Teles AR, da Silva PG, Martins D, Guyot JP, et al. Twelve years of scientific production on Medline by Latin American spine surgeons. **PLoS One 2014**; 9(2): e87945.
60. Lynn J, McKethan A, Jha AK. Value-based payments require valuing what matters to patients. **JAMA 2015**; 314(14): 1445–1446.

How to cite this article: Teles AR, Khoshhal KI, Falavigna A. Why and how should we measure outcomes in spine surgery?. Journal of Taibah University Medical Sciences 2016;11:91–97.