

Assessment of the minimum clinically important difference in neurological function and quality of life after surgery in cervical spondylotic myelopathy patients: a prospective cohort study

Feifei Zhou¹ · Yilong Zhang¹ · Yu Sun¹ · Fengshan Zhang¹ · Shengfa Pan¹ · Zhongjun Liu¹

Received: 29 January 2015/Revised: 22 August 2015/Accepted: 22 August 2015/Published online: 1 September 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract

Purpose The purpose of this study was to evaluate the application of the minimum clinically important difference (MCID) concept to postoperative clinical results by using a prospective cohort study in Chinese patients with cervical spondylotic myelopathy (CSM).

Methods The sample consisted of 113 patients who underwent surgical treatment for CSM in our hospital between February 2008 and November 2012. The preoperative and 1-year postoperative modified Japanese Orthopaedic Association (mJOA) scores, mJOA score recovery rate, physical component summary (PCS) and mental component summary (MCS) of the Short Form 36 were collected. The MCID of each outcome measurement was calculated by four approaches including average change, minimum detectable change, change difference and receiver operating characteristic curve. The responsiveness of each measurement was then analyzed.

Results The patients presented a statistically significant improvement ($p < 0.01$) postoperatively in mJOA, PCS, and MCS. The MCID calculated by four approaches varied from 4.09 to 9.62 for the PCS, 3.11 to 7.41 for the MCS, 1.25 to 3.07 for mJOA score, and 31.37 to 44.02 % for mJOA recovery rate. In addition, the improvement of the mJOA score owned the highest responsiveness of the four outcome measurements.

Conclusions The threshold value of the MCID was determined by the choice of the assessment approach. In addition, the recovery rate of the mJOA score appeared to be the most valid and responsive measure of effectiveness of surgery in CSM patients.

Keywords Cervical spondylotic myelopathy · Cervical spine · Minimum clinically important difference · Modified Japanese Orthopaedic Association score · Short Form 36

Introduction

Cervical spondylotic myelopathy (CSM) is one of the most prevalent observed neurological disorders in the geriatric population [1, 2]. In the past, evaluations of the postoperative curative effects of the procedure were primarily based on improvements in neurological dysfunction, such as that defined by the modified Japanese Orthopaedic Association (mJOA) score [3]. In recent years, more attention has gradually been paid to the subjective feelings of patients using, for example, the Short Form 36 (SF-36) of the Medical Outcomes Study [4]. A comprehensive clinical assessment should take both disease-specific evaluations and general health evaluations into consideration.

The majority of studies that attempt to quantify the efficacy of a therapeutic intervention report changes in group means before and after treatment. However, group means cannot be readily used in clinical practice to interpret changes on an individual bias [5]. Therefore, the concept of the “minimum clinically important difference” (MCID) has been presented as an alternative tool with which to quantify clinically significant patient improvements due to a therapeutic intervention. The MCID is defined as the smallest change that is meaningful to

F. Zhou and Y. Zhang contributed equally to this work.

✉ Yu Sun
sunyuor@163.com

¹ Department of Orthopedics, Peking University Third Hospital, Beijing 100191, China

patients and is considered the threshold needed to achieve treatment efficacy [6]. A MCID value in a given study that exceeds a threshold value indicates that a clinically significant change was achieved, which may validate a “decision to treat [7]”.

The purpose of this study was to calculate the MCID for the SF-36 and mJOA score and to evaluate the responsiveness of each outcome measurements in CSM patients who underwent surgery.

Methods

Patient sample

This prospective cohort study was approved by the Institutional Ethics Committee of our Hospital. The study was performed in 141 consecutive patients with CSM treated by the same group of spine surgeons in our hospital between February 2008 and November 2012. The diagnosis of each patient was made based on an integrative analysis of the patient’s symptoms and physical and radiological examinations, including preoperative cervical anteroposterior, lateral, and flexion–extension radiographs and magnetic resonance imaging (MRI) or computed tomography (CT) scans. The diagnosis was either CSM or ossification of the posterior longitudinal ligament (OPLL) patients with myelopathy. All of the patients were refractory to conservative treatment. Patients with trauma, infection, or serious complications due to hypertension, coronary heart disease, or diabetes, as well as patients who had these conditions prior to surgery, regardless of whether the condition(s) was under control, were excluded. Patients who did not complete the outcome questionnaires prior to surgery or at the 1-year follow-up were also excluded. A total of 113 (80.1 %, 113/141) consecutive patients was enrolled. In each case, the patient completed the questionnaires with the assistance of the same specially trained medical personnel. The follow-ups were scheduled at an outpatient clinic.

Outcome assessment

The SF-36 is a 36-item self-administered short-form health status survey that was developed in the Medical Outcomes Study. The SF-36 primarily evaluates a patient’s social and physical function, general health, vitality, and body pain. Two composite scores can be obtained, namely, a physical composite summary (PCS) score and a mental composite summary (MCS) score. These values were used in the general health outcome assessments [8].

The modified Japanese Orthopaedic Association (mJOA) score for cervical myelopathy [0 (lowest) to 17 points (highest)] was recorded before surgery and at 1 year after

surgery. This score assesses neurological function based on three categories: motor function in the upper and lower limbs, sensory function in the upper and lower limbs, and trunk and bladder function. The Hirabayashi recovery rates were calculated using the formula established by Hirabayashi: $(\text{postoperative JOA score} - \text{preoperative JOA score}) / (17 - \text{preoperative JOA score}) \times 100 (\%)$. These values were used for the disease-specific outcome evaluation [9].

Calculating approaches for MCID

There is no “gold standard” methodology for estimating the MCID. The majority of methods fall within the following two categories, distribution-based methods and anchor-based methods, including the following four approaches [10]: (1) “mean change”, an MCID value that correlates with the average change in the patient cohort that exhibits small variations [11]; (2) “change difference”, the difference in the average change score between responders and nonresponders [12]; (3) “minimum detectable change” (MDC), the smallest value that is greater than the measurement error within a 95 % confidence interval (CI) [13]; and (4) “receiver operating characteristic (ROC) curve”, a sensitivity- and specificity-based approach for the calculation of the MCID.

Anchors

The health transition item (HTI) of the SF-36 was used as the anchor for derivations of anchor-based MCID calculations; the HTI is a well-studied anchor and has been used extensively in many previous studies [10, 11]. The HTI assesses how a patient feels at the time of the questionnaire completion compared with 1 year previously. The HTI is considered an appropriate independent anchor because it is not used in the scoring of the MCS or PCS of the SF-36.

The responsiveness of the MCID of each outcome measurement

When used in the arena of the MCID, sensitivity is defined as the proportion of patients who report improvements based on the external criterion and who have a patient-reported outcome (PRO) score greater than the MCID threshold value. Specificity is defined as the proportion of patients who do not report improvements based on the external criterion and who have a PRO score less than the MCID threshold value. The ROC curve-derived MCID was the change score with equal sensitivity and specificity. The accuracy of the ROC curve is evaluated using the calculated area under the curve (AUC). AUC in the range of 0.90–1.00 are considered excellent, 0.80–0.90 are considered good, 0.60–0.80 are considered fair, and 0.50–0.60 are

considered to indicate failure [14]. Thus, to evaluate the responsiveness of the MCID of each outcome assessment, we calculated AUC of the ROC curve and correlations between responses to the anchor.

Statistical analysis

SPSS version 17.0 software (SPSS Inc., USA) was used for the statistical analysis. The preoperative and 1-year postoperative scores were compared using the Mann–Whitney *U* test. Correlation coefficients were calculated to evaluate the associations between the answers to the anchor question and the outcome scores (Spearman coefficients). *P* values of less than 0.05 were considered statistically significant.

Results

Patient group

A total of 141 consecutive patients completed the preoperative outcome measures. Of these patients, 113 (80.1 %, 113/141) completed all of the outcome measures 1 year

after surgery. The mean age of the patients at baseline was 57.6 ± 11.1 years. Of the patients, 39.8 % (45/113) were female and 60.2 % (68/113) were male. The surgical information of patients is shown in Table 1.

MCID threshold values for the outcome measurements

All outcome measurements appeared a significant improvement after surgery with 1-year follow-up (Table 2). The comparison of different anchor- and distribution-based approaches yielded a wide range of MCID threshold values for each outcome measure (Table 3). These values varied from 4.09 to 9.62 for the PCS, 3.11 to 7.41 for the MCS, 1.25 to 3.07 for the mJOA score, and 31.37 to 44.02 % for the recovery rate of the mJOA score.

The responsiveness of the MCID of each outcome measurement

To determine which outcome measure was the most valid and responsive measure of therapeutic effectiveness in CSM patients undergoing surgery, we used ROC curves to

Table 1 Demographics and types of surgery

Age (years)	Mean 57.6	Range 32–90	
Gender	Male = 45 (39.8 %)	Female = 68 (60.2 %)	
Treatment	Posterior approach	Laminoplasty	49 (43.4 %)
		ACDF	47 (41.6 %)
	Anterior approach	Artificial disc replacement	11 (9.7 %)
		ACCF	6 (5.3 %)

ACDF anterior cervical decompression and fusion, *ACCF* anterior cervical corpectomy and fusion

Table 2 Patient-reported outcomes (SF-36 PCS, SF-36 MCS, mJOA score, recovery rate of the mJOA score) at baseline and 1 year after surgery

Outcome score	Preoperative scores, mean (SD)	1-year postoperative score, mean (SD)	<i>p</i> values
SF-36 PCS	37.85 (6.58)	43.29 (7.27)	<0.01
SF-36 MCS	39.53 (5.54)	43.65 (7.06)	<0.01
mJOA	12.45 (2.51)	15.15 (1.50)	<0.01
Recovery rate	–	60.88 % (37.73 %)	–

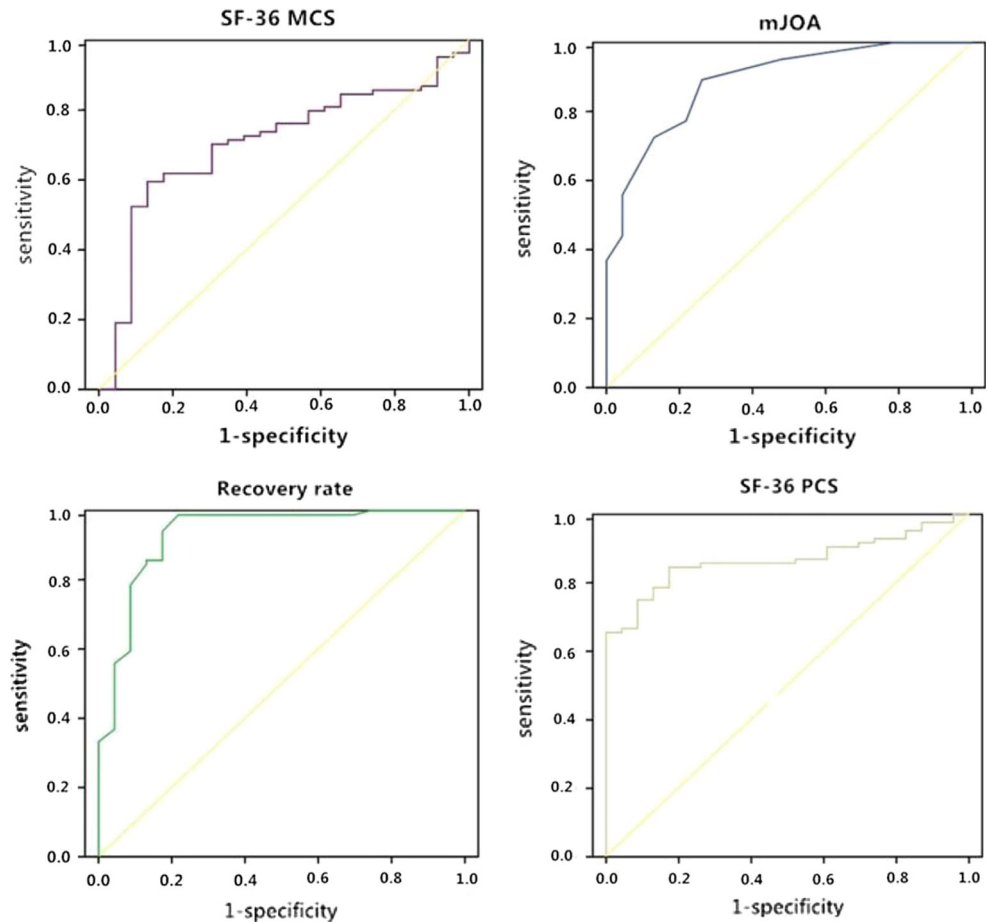
SF-36 PCS SF-36 physical component summary score, *SF-36 MCS* SF-36 mental component summary score, *mJOA score* modified Japanese Orthopaedic Association score

Table 3 MCID threshold values for PCS, MCS, mJOA score, and recovery rate of the mJOA score

MCID calculation method	Outcome measure			
	PCS	MCS	mJOA score	Recovery rate (%)
Average change	5.44	3.11	2.70	–
MDC (95 % CI)	5.16	3.43	2.08	31.37
Change difference	9.62	7.41	3.07	44.02
ROC curve derived	4.09	3.91	1.25	31.67

SF-36 PCS SF-36 physical component summary score, *SF-36 MCS* SF-36 mental component summary score, *mJOA score* modified Japanese Orthopaedic Association score, *MCID* minimum clinically important difference, *MDC* minimum detectable change, *ROC curve* receiver operating characteristic curve

Fig. 1 Receiver operating characteristic (ROC) curves for each outcome assessment. The AUCs for the mJOA score, the mJOA score recovery rate, PCS, and MCS were, 0.89, 0.93, 0.87, and 0.71, respectively. *ROC curve* receiver operating characteristic curve, *AUC* area under the curve



compare the four outcome measures (mJOA score, mJOA score recovery rate, PCS, and MCS of SF-36) assessed in our study. Besides, the association between the responses to the anchor and the change in outcome measurements was also calculated.

The AUC varied from 0.71 to 0.93, indicating that the ROC curve exhibited suitable accuracy discriminating between responders and nonresponders (Fig. 1). The AUCs for the mJOA score, the mJOA score recovery rate, PCS, and MCS were, 0.89, 0.93, 0.87, and 0.71, respectively. The recovery rate appeared to be the most accurate discriminator of meaningful effectiveness (AUC of 0.93) and the most responsive to postoperative improvement.

The greatest correlations were between the responses to the anchor and the recovery rate of the mJOA score ($r = 0.796$, $p < 0.01$). The correlations were nearly identical between responses to the anchor and the mJOA score ($r = 0.537$, $p < 0.01$) and the PCS ($r = 0.532$, $p < 0.01$), whereas the correlation between the response to the anchor and the MCS was $r = 0.332$ ($p < 0.01$).

Discussion

There are literatures supporting the use of both a generic and a disease-specific questionnaire because these questionnaires evaluate different, although complementary, aspects of patient outcomes [15, 16]. A subjective outcome questionnaire could be used to evaluate patient pain, dysfunction, and general health status; however, the specific value of each score does not reflect actual clinical significance and offers little assistance in guiding or assessing clinical practices. Furthermore, the patients' subjective experiences could be quite different even when the patient-reported outcome scores are the same, which could cause difficulty in evaluating the efficacy of clinical treatment with specific scores.

In this study, a disease-specific questionnaire, the mJOA, and the recovery rate of the mJOA score were used to evaluate patients' neurological function improvements, whereas the PCS and MCS in SF-36 were used to evaluate general condition improvements. The changes in the mJOA score, mJOA score improvement rate, MCID on PCS, and

MCS of SF-36 were used to evaluate clinical significance. The latter calculation facilitated the detection of the smallest change in an outcome measure that reflects clinically meaningful improvements in patients with CSM. Clinical treatment was deemed significantly effective when its score reached the MCID threshold value.

Three previous studies compared different MCID calculation methods in patients with cervical disorders. Carreon [5] described MCID values in a heterogeneous population undergoing both anterior and posterior cervical spine fusion; however, that study did not include the MCS of the SF-36. Parker [17] assessed anchor-based approaches in patients with cervical radiculopathy undergoing ACDF; however, there was only a 3-months follow-up. More recently, Auffinger [7] described the MCID of ACDF based on a long-term follow-up of only 30 patients. To the best of our knowledge, there are no previous studies that determined the MCID, neurological function, and patients' subjective evaluations with a relatively long follow-up period.

MCID threshold is manifested as an interval value for different choice of MCID calculation method leads to different result; however, there is no consensus regarding the best method. Several previous studies, such as Copay [18] and Parker's [19] research, suggested use of the MDC as the optimal method for determining the MCID threshold in patients undergoing spine surgeries.

To evaluate which assessment is the most valid and responsive measure of the therapeutic effectiveness of CSM, we calculated the ROC curve and correlations between responses to the anchor. The ROC curve attempts to identify the threshold for a PRO score while maintaining the greatest sensitivity and specificity possible and attempts to quantify personal bias. The area under the ROC curve corresponding to the feasibility of using the change in the PCS score to predict improvement was 0.87. This value could be interpreted as follows: if one takes two individuals at random, namely, one with and without improvements according to the global effectiveness rating, the probability is 87 % that the first individual would have a higher change score than the second individual. Similarly, the AUCs of the MCS, mJOA score, and ROC curve of the mJOA score improvement rate were 0.71, 0.89, and 0.93, respectively, indicating that the largest variation in neurological function corresponded to the mJOA score improvement. The correlation coefficient between the mJOA score improvement and HTI score reached 0.796 ($r = 0.796, p < 0.01$), which is larger than the correlation coefficient between HTI and the PCS ($r = 0.532, p < 0.01$), MCS ($r = 0.332, p < 0.01$), and mJOA ($r = 0.537, p < 0.01$). The mJOA score improvement rate is the measure that appeared to be the most accurate indicator of meaningful effectiveness

(AUC of 0.93) and most responsive to postoperative improvement ($r = 0.796, p < 0.01$).

The large value of the AUC of the mJOA improvement rate and the strong correlation between the mJOA improvement rate and the HTI suggest that the patients' neurological function was improved together with their health status 1 year after the surgery. Patients with CSM pay more attention to their neurological functional improvement than their physiological function and mental state. In this study, it appeared that Chinese patients were generally keen on discussing the improvement of previous symptoms due to CSM during subsequent visits. In addition, pathogenic factors and the natural course of CSM add complexity and unpredictability to the postoperative outcomes. Surgical treatment for spinal disorders is aimed at improving symptoms, quality of life, work ability, and social activity [20]. We could not ignore patients' mental health even though the correlation coefficient between the MCS and HTI was low ($r = 0.332, p < 0.01$). The result that the improvement in the MCS was less than that of the PCS indicates that patients' mental health may play a more important role than expected. A large-scale study is needed to further investigate the postoperative long-term mental health states of patients with CSM.

There are several limitations to this study. First, the patients were not subdivided into subgroups according to surgical method because of the limited number of cases. For example, the prevalence of OPLL ranged from 0.4 to 3.0 % in Asian countries, and OPLL patients with myelopathy might be a special group of CSM patients [21]. Further studies are still needed to be carried out. Second, the MCID values are specific to the populations studied and maybe different in other populations depending on their specific characteristics, such as gender and work status. Finally, the anchor utilized in this study is not comprehensive because other anchors, such as the surgeon rating and visual analogue scale (VAS) score, were not included. Further multicenter studies with larger datasets and perhaps longer follow-up times are encouraged to examine whether the MCIDs differ in different indications and subgroups of surgically operated patients.

Conclusion

All outcome assessments for CSM patients improved significantly 1-year after surgery. The threshold value of the MCID for the surgical treatment of patients with CSM was determined to a great extent by the choice of the assessment method. In addition, the recovery rate of the mJOA score appeared to be the most valid and responsive measure of effectiveness in CSM patients undergoing surgery.

Compliance with ethical standards**Conflict of interest** None.**References**

- Chang V, Holly LT (2013) Controversies in the management of cervical spondylotic myelopathy. *J Neurosurg Sci* 57:241–252
- Fehlings MG, Arvin B (2009) Surgical management of cervical degenerative disease: the evidence related to indications, impact, and outcome. *J Neurosurg Spine* 11:97–100. doi:[10.3171/2009.5.SPINE09210](https://doi.org/10.3171/2009.5.SPINE09210)
- Grosso MJ, Hwang R, Mroz T, Benzel E, Steinmetz MP (2013) Relationship between degree of focal kyphosis correction and neurological outcomes for patients undergoing cervical deformity correction surgery. *J Neurosurg Spine* 18:537–544. doi:[10.3171/2013.2.SPINE12525](https://doi.org/10.3171/2013.2.SPINE12525)
- Singh A, Gnanalingham K, Casey A, Crockard A (2006) 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)* 31:639–643. doi:[10.1097/01.brs.0000202744.48633.44](https://doi.org/10.1097/01.brs.0000202744.48633.44)
- Carreon LY, Glassman SD, Campbell MJ, Anderson PA (2010) 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* 10:469–474. doi:[10.1016/j.spinee.2010.02.007](https://doi.org/10.1016/j.spinee.2010.02.007)
- Lauche R, Langhorst J, Dobos GJ, Cramer H (2013) Clinically meaningful differences in pain, disability and quality of life for chronic nonspecific neck pain—a reanalysis of four randomized controlled trials of cupping therapy. *Complement Ther Med* 21:342–347. doi:[10.1016/j.ctim.2013.04.005](https://doi.org/10.1016/j.ctim.2013.04.005)
- Auffinger B, Lam S, Shen J, Thaci B, Roitberg BZ (2013) Usefulness of minimum clinically important difference for assessing patients with subaxial degenerative cervical spine disease: statistical versus substantial clinical benefit. *Acta Neurochir (Wien)* 155:2345–2354. doi:[10.1007/s00701-013-1909-4](https://doi.org/10.1007/s00701-013-1909-4) (**discussion 2355**)
- King JT Jr, McGinnis KA, Roberts MS (2003) Quality of life assessment with the medical outcomes study Short Form 36 among patients with cervical spondylotic myelopathy. *Neurosurgery* 52:113–120 (**discussion 121**)
- Machino M, Yukawa Y, Hida T, Ito K, Nakashima H, Kanbara S, Morita D, Kato F (2012) The prevalence of pre- and postoperative symptoms in patients with cervical spondylotic myelopathy treated by cervical laminoplasty. *Spine (Phila Pa 1976)* 37:E1383–E1388. doi:[10.1097/BRS.0b013e3182684c68](https://doi.org/10.1097/BRS.0b013e3182684c68)
- Brozek JL, Guyatt GH, Schunemann HJ (2006) How a well-grounded minimal important difference can enhance transparency of labelling claims and improve interpretation of a patient reported outcome measure. *Health Qual Life Outcomes* 4:69. doi:[10.1186/1477-7525-4-69](https://doi.org/10.1186/1477-7525-4-69)
- Juniper EF, Guyatt GH, Willan A, Griffith LE (1994) Determining a minimal important change in a disease-specific Quality of Life Questionnaire. *J Clin Epidemiol* 47:81–87
- Hagg O, Fritzell P, Nordwall A, Swedish Lumbar Spine Study G (2003) The clinical importance of changes in outcome scores after treatment for chronic low back pain. *Eur Spine J* 12:12–20. doi:[10.1007/s00586-002-0464-0](https://doi.org/10.1007/s00586-002-0464-0)
- Wells G, Beaton D, Shea B, Boers M, Simon L, Strand V, Brooks P, Tugwell P (2001) Minimal clinically important differences: review of methods. *J Rheumatol* 28:406–412
- van der Roer N, Ostelo RW, Bekkering GE, van Tulder MW, de Vet HC (2006) Minimal clinically important change for pain intensity, functional status, and general health status in patients with nonspecific low back pain. *Spine (Phila Pa 1976)* 31:578–582. doi:[10.1097/01.brs.0000201293.57439.47](https://doi.org/10.1097/01.brs.0000201293.57439.47)
- Dattani R, Slobogean GP, O'Brien PJ, Broekhuysen HM, Blachut PA, Guy P, Lefaivre KA (2013) Psychometric analysis of measuring functional outcomes in tibial plateau fractures using the Short Form 36 (SF-36), Short Musculoskeletal Function Assessment (SMFA) and the Western Ontario McMaster Osteoarthritis (WOMAC) questionnaires. *Injury* 44:825–829. doi:[10.1016/j.injury.2012.10.020](https://doi.org/10.1016/j.injury.2012.10.020)
- Bombardier C, Melfi CA, Paul J, Green R, Hawker G, Wright J, Coyte P (1995) Comparison of a generic and a disease-specific measure of pain and physical function after knee replacement surgery. *Med Care* 33:AS131–AS144
- Parker SL, Godil SS, Shau DN, Mendenhall SK, McGirt MJ (2013) 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* 18:154–160. doi:[10.3171/2012.10.SPINE12312](https://doi.org/10.3171/2012.10.SPINE12312)
- Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY (2008) 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* 8:968–974. doi:[10.1016/j.spinee.2007.11.006](https://doi.org/10.1016/j.spinee.2007.11.006)
- Parker SL, Mendenhall SK, Shau DN, Adogwa O, Anderson WN, Devin CJ, McGirt MJ (2012) 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* 16:471–478. doi:[10.3171/2012.1.SPINE11842](https://doi.org/10.3171/2012.1.SPINE11842)
- Toledano M, Bartleson JD (2013) Cervical spondylotic myelopathy. *Neurol Clin* 31:287–305. doi:[10.1016/j.ncl.2012.09.003](https://doi.org/10.1016/j.ncl.2012.09.003)
- Matsunaga S, Sakou T (2012) Ossification of the posterior longitudinal ligament of the cervical spine: etiology and nature history. *Spine (Phila Pa 1976)* 37:E309–E314. doi:[10.1097/BRS.0b013e318241ad33](https://doi.org/10.1097/BRS.0b013e318241ad33)