Minimum Clinically Important Differences in the Cervical Spine Outcomes Questionnaire

Results from a National Multicenter Study of Patients Treated with Anterior Cervical Decompression and Arthrodesis

Richard L. Skolasky, ScD, Todd J. Albert, MD, Anica M. Maggard, BS, and Lee H. Riley III, MD

Background: The minimum clinically important difference is a clinically relevant threshold of improvement. A substantial clinical benefit is a threshold of change that correlates with clinically important improvement. The Cervical Spine Outcomes Questionnaire is a disease-specific, patient-reported outcomes instrument that was developed to be sensitive to changes associated with surgical treatment for degenerative cervical disc disease. To determine thresholds for change in these domain scores that are important from the patient's perspective, we estimated the minimum clinically important difference and substantial clinical benefit values for this questionnaire's domain scores.

Methods: We evaluated 252 patients from the Cervical Spine Research Society Outcomes Study at their six-month follow-up visits after anterior cervical spine decompression and arthrodesis. Using a receiver operating characteristics curve, with the health transition item of the Short Form-36 as an anchor, we determined that the minimum clinically important difference (the value that maximized sensitivity and specificity to differentiate the ''somewhat better'' and ''much better'' responses from others) and the substantial clinical benefit (the value that maximized sensitivity and specificity to differentiate the ''much better'' response from others) for our questionnaire's domain scores. Responses were scaled between 0 and 1 point; higher scores denoted more severe impairment. Patient and clinical characteristics were tested to determine their influence on score changes.

Results: The minimum clinically important difference ranged from 0.13 point (for functional disability) to 0.24 point (for arm/shoulder pain). The substantial clinical benefit score ranged from 0.20 point (for functional disability or physical symptoms other than pain) to 0.30 point (for neck or arm/shoulder pain). Age, sex, and duration of current symptoms were not associated with change in our questionnaire's domain scores.

Conclusions: A 0.13-point change in the functional disability domain score indicated a clinically important difference in a self-reported outcome after anterior cervical spine surgery. A 0.30-point change in neck pain after surgery indicated a clinically important clinical benefit. This information, coupled with previous reports of the psychometric stability of the Cervical Spine Outcomes Questionnaire, should increase the clinical utility of this patient-reported outcomes instrument.

The use of patient-reported outcomes instruments to gauge recovery after treatment and satisfaction with care has become more common. Reports in the medical literature often cite mean differences between pretreatment and posttreatment scores or between two different treatment groups. The reliance on mean differences does not provide the reader with the ability to incorporate the use of patient-reported outcomes into clinical practice.

The minimum clinically important difference has been defined as the smallest change that is meaningful to patients[1,2]. This difference has been estimated for a number of commonly used patient-reported outcomes measures, including the Short Form-36 (SF-36) and the Neck Disability Index (NDI). The SF-36 is a generic health status measure that was developed to provide an estimate of an individual's physical and mental health. It has been estimated that the minimum clinically important difference in SF-36 physical health component scores can range from 3 to 6 points[3,4]. The NDI is a disease-specific, patient-reported outcomes instrument that is used to assess an individual's level of disability. The minimum clinically

Disclosure: One or more of the authors received payments or services, either directly or indirectly (i.e., via his or her institution), from a third party in support of an aspect of this work. In addition, one or more of the authors, or his or her institution, has had a financial relationship, in the thirty-six months prior to submission of this work, with an entity in the biomedical arena that could be perceived to influence or have the potential to influence what is written in this work. Also, one or more of the authors has had another relationship, or has engaged in another activity, that could be perceived to influence or have the potential to influence what is written in this work. The complete **Disclosures of Potential Conflicts of Interest** submitted by the authors of this work are available with the online version of this article at jbjs.org.

TABLE I Baseline and Six-Month Patient-Reported Outcomes Scores

Patient Reported Outcomes Instrument	Baseline Score* *(points)*	Six-Month Score* *(points)*	Change in Score* *(points)*
NDI	47.88 ± 15.57	35.43 ± 17.23	–12.38 ± 15.68
SF-36			
Physical health	50.18 ± 10.55	57.29 ± 11.64	7.11 ± 9.37
Mental health	53.73 ± 10.93	59.48 ± 11.52	5.75 ± 10.90
CSOQ			
Neck pain	0.53 ± 0.24	0.25 ± 0.24	–0.28 ± 0.24
Arm/shoulder pain	0.56 ± 0.20	0.26 ± 0.25	–0.30 ± 0.26
Functional disability	0.41 ± 0.20	0.20 ± 0.21	–0.21 ± 0.22
Psychologic distress	0.53 ± 0.21	0.34 ± 0.21	–0.17 ± 0.21
Physical symptoms other than pain	0.60 ± 0.30	0.43 ± 0.24	–0.17 ± 0.27

*The values are given as the mean ± standard deviation.

important difference in NDI scores has been estimated to be 7.5 points[5,6].

The Cervical Spine Outcomes Questionnaire (CSOQ) (see Appendix) is a multidimensional patient-reported outcomes instrument that was developed to specifically assess outcomes after anterior cervical spine decompression and arthrodesis in a Cervical Spine Research Society-funded outcomes study. The CSOQ provides estimates of neck and arm/shoulder pain, physical symptoms other than pain, functional disability, psychologic distress, and health-care utilization. Previous reports have shown the measure to have good psychometric properties[7] and utility for tracking patient recovery[8]. We are not aware of any study in which the investigators estimated the minimum clinically important difference for the domain scores of the CSOQ.

Our goal was to determine the CSOQ scores that represent a minimum clinically important difference and a substantial clinical benefit for patients treated with anterior cervical spine decompression and arthrodesis for cervical degenerative disc disease. We also tested the influence of patient and clinical characteristics on change in CSOQ domain scores.

Materials and Methods

This study was approved by our institutional review board.

Patient Population

This study is a retrospective review of data that were prospectively collected from a cohort of patients undergoing anterior cervical spine decompression and arthrodesis at one of twenty-three national sites enrolled in the Cervical Spine Research Society Outcomes Study. In this study, operating surgeons were orthopaedists or neurosurgeons at centers that specialized in spine care. Inclusion criteria for the Cervical Spine Research Society Outcomes Study were the presence of cervical radiculopathy or myelopathy, use of an anterior surgical approach, adequate mental capacity to provide informed consent, patient age of more than eighteen years, complete enrollment information, and the ability to return for follow-up assessment. An additional criterion for inclusion in the current study was that the individual had to have provided complete information on the SF-36, the NDI, and the CSOQ at baseline and the six-month postoperative assessment. Of the 622 patients who had undergone surgery, 370 were excluded for the following reasons: posterior cervical spine surgery, 126; missing enrollment forms (including completion of the NDI, SF-36, and CSOQ), ninety-eight; missing six-month forms, eighty-five; absence of cervical radiculopathy or myelopathy, thirty-six; and requested withdrawal from the study, twenty-five. The remaining 252 individuals (40.5%) met all of the criteria and formed our study group.

This sample of participants tended to be middle-aged (mean age at surgery, 49.8 ± 10.4 years), white (87%), and well-educated (median education, fourteen years; interquartile range, eleven to fifteen years). Both sexes were equally represented (male, 55.6%); 142 (56.3%) of the participants reported smoking cigarettes at some point in the past, with fifty-one (20.2%) reporting current tobacco use. Eight participants (3.2%) were receiving unemployment benefits, thirty-four (13.5%) reported receiving Workers' Compensation payments, and twenty-nine (11.5%) were receiving Social Security disability benefits. Of the 252 participants, 103 (40.9%) underwent single-level cervical spine surgery; ninety-seven (38.5%), two-level surgery; and the remaining fifty-two (20.6%), surgery involving three or more levels. Radiculopathy was the primary problem in 153 individuals (60.7%); myelopathy, in sixty-one (24.2%); and axial pain, in the remaining thirty-eight (15.1%).

Assessment Instruments

Baseline and postoperative evaluations included use of the NDI, the SF-36, and the CSOQ.

The NDI is a disease-specific instrument that is used to assess the impact of spinal disorders on ten aspects of daily living[9]. The NDI has excellent retest reliability (Pearson $r > 0.80$) and validity (moderately high correlations with the McGill Pain Questionnaire and a visual analog scale for pain). An expert panel convened by the journal *Spine* recommended use of this instrument because of its good psychometric properties[5].

The SF-36 is a rater-administered survey that is used to assess the severity of an individual's functional disability related to his or her physical or emotional health[10]. This instrument is a reliable measure of disability and general health status and has been used extensively in similar populations. The SF-36 has excellent retest reliability (Pearson $r = 0.93$)[11].

We used an anchor-based method that corresponded to a minimum clinically important difference and a substantial clinical benefit[12]. The health transition item of the SF-36, not used in calculating the SF-36 aggregate health scores, was used as the anchor. This item asks respondents: "Compared with one year ago today, how would you rate your health in general now?" Answer

TABLE II Change in Patient-Reported Outcomes Scores, Stratified by Response to Anchor Question

		Mean Patient-Reported Outcomes Scores* *(points)*		
			SF-36	
Anchor Response	No.	NDI	Physical Component	Mental Component
Much worse	8	0.93 ± 15.04	–2.78 ± 8.85	–3.04 ± 11.42
Somewhat worse	20	–8.49 ± 17.16	4.03 ± 9.79	2.52 ± 11.11
About the same	66	–9.61 ± 18.99	4.56 ± 8.80	3.57 ± 10.24
Somewhat better	74	–11.72 ± 11.80	5.95 ± 6.50	5.25 ± 8.96
Much better	84	–19.18 ± 14.37	11.76 ± 9.92	10.24 ± 11.06
P value†		<0.001	<0.001	<0.001

*The values are given as the mean ± standard deviation. †Represents significance demonstrated by across-group comparisons with the health transition item strata (anchor responses), with use of analysis of variance.

choices include "much worse," "somewhat worse," "about the same," "somewhat better," and "much better." Participants responding "somewhat better" and "much better" were considered to be experiencing a clinically important improvement in self-reported health outcome. Participants responding "much better" were considered to be experiencing a clinically important improvement in self-reported health outcome.

The CSOQ measures patient-reported outcomes across several broad factors, including pain severity (in the neck and arms/shoulders), functional disability, psychologic distress, physical symptoms other than pain, and health-care utilization (Table I)[8]. The CSOQ is a valid and reliable outcomes instrument (Pearson r > 0.80) for patients undergoing cervical spine surgery[7].

Statistical Analysis

To calculate the minimum clinically important difference, we used receiver operating characteristics curve analysis to plot the sensitivity and specificity for each patient-reported outcome. We chose the optimal cutoff for the receiver operating characteristics curve to maximize the sensitivity and specificity[12]. This value represented the minimum clinically important difference. We also used receiver operating characteristics curve analysis and chose the optimal cutoff to calculate the substantial clinical benefit. We used the minimum clinically important difference and the substantial clinical benefit to classify each individual and compared those classifications with the anchor stratification.

In addition, the standard error measure was calculated to represent the precision of the measurement for each patient-reported outcome instrument in this population. The standard error measure is calculated as the product of the sample standard deviation and the square root of one minus the test-retest reliability of the instrument. Published test-retest reliability estimates were used for each scale: 0.8 for the NDI[9], 0.93 for the SF-36[4], and 0.90 for the CSOQ[7].

A secondary analysis was performed to determine the influence of patient and clinical factors on change in CSOQ domain scores. The patient factors were age and sex, and the clinical factor was the duration of current symptoms. We used a t test to compare the mean changes in the CSOQ domain scores among patients stratified by sex. Spearman correlation coefficients were estimated to examine the relationship between the change in the CSOQ domain scores and age and duration of current symptoms.

Significance was set at $p < 0.05$.

Source of Funding

Data for this report were collected in the Cervical Spine Research Society Outcomes Study, which is funded through the Cervical Spine Research Society with contributions from Aesculap, DePuy Spine, Medtronic Sofamor Danek, Stryker Spine (Osteonics), and Synthes Spine.

Results

There was significant improvement, from the preoperative baseline to the six-month postoperative assessment, in the CSOQ domain scores for neck and arm/shoulder pain, functional disability, psychologic distress, and physical symptoms other than pain ($p < 0.001$ for all) (Table I). Similarly, there was significant improvement in the NDI scores and in the physical and mental health component scores of the SF-36 ($p < 0.001$ for all) (Table I).

Seventy-four patients (29.4%) indicated "somewhat better" and eighty-four (33.3%) indicated "much better" in response to the health transition item of the SF-36 (Table II). There were significant differences across health transition item strata with regard to the changes, between the baseline and six-month assessments, in the CSOQ domain scores for neck and arm/shoulder pain ($p < 0.001$ and $p = 0.002$, respectively), functional disability ($p = 0.001$), psychologic distress ($p = 0.003$), and physical symptoms other than pain ($p = 0.012$). Similarly, there were significant differences across health transition item strata with respect to the changes in the NDI scores ($p < 0.001$) and the SF-36 physical and mental health component scores ($p < 0.001$ for each). On each of these patient-reported outcomes instruments, individuals who reported "somewhat better" and "much better" health experienced greater changes between the baseline and six-month assessments than did those reporting "somewhat worse" and "much worse" health.

Of the 252 individuals, 158 (62.7%) reported achieving a minimum clinical benefit. The minimum clinically important differences in the CSOQ domain scores, as determined by patients' responses to the anchor question, were: neck pain, 0.20 point; arm/shoulder pain, 0.24 point; functional disability, 0.13 point; psychologic distress, 0.17 point; and physical symptoms other than pain, 0.20 point (Table III). The minimum clinically important differences of other outcome measures were: NDI, 11.3 point; SF-36 physical component score, 6.5 point; and SF-36 mental component score, 5.0 point (Fig. 1).

TABLE II (*continued*)

Mean Patient-Reported Outcomes Scores* *(points)*				
CSOQ				
Neck Pain	Arm/Shoulder Pain	Functional Disability	Psychologic Distress	Physical Symptoms Other Than Pain
–0.04 ± 0.20	–0.09 ± 0.20	0.02 ± 0.16	–0.09 ± 0.15	–0.08 ± 0.18
–0.15 ± 0.27	–0.28 ± 0.25	–0.14 ± 0.21	–0.12 ± 0.24	–0.05 ± 0.23
–0.22 ± 0.21	–0.25 ± 0.27	–0.19 ± 0.22	–0.16 ± 0.23	–0.12 ± 0.24
–0.28 ± 0.22	–0.28 ± 0.28	–0.21 ± 0.21	–0.19 ± 0.19	–0.19 ± 0.28
–0.38 ± 0.23	–0.39 ± 0.24	–0.27 ± 0.20	–0.27 ± 0.20	–0.23 ± 0.27
<0.001	0.002	0.001	0.003	0.012

Of the 252 individuals, eighty-four (33.3%) reported achieving substantial clinical benefit. The change in scores representing substantial clinical benefit in the CSOQ domains were: neck pain, 0.30 point; arm/shoulder pain, 0.30 point; functional disability, 0.20 point; psychologic distress, 0.22 point; and physical symptoms other than pain, 0.20 point (Table III). The change in scores representing substantial clinical benefit for the other outcome measures were: NDI, 13.0 points; SF-36 physical component score, 7.8 points; and SF-36 mental component score, 6.6 points (Table III, Fig. 2).

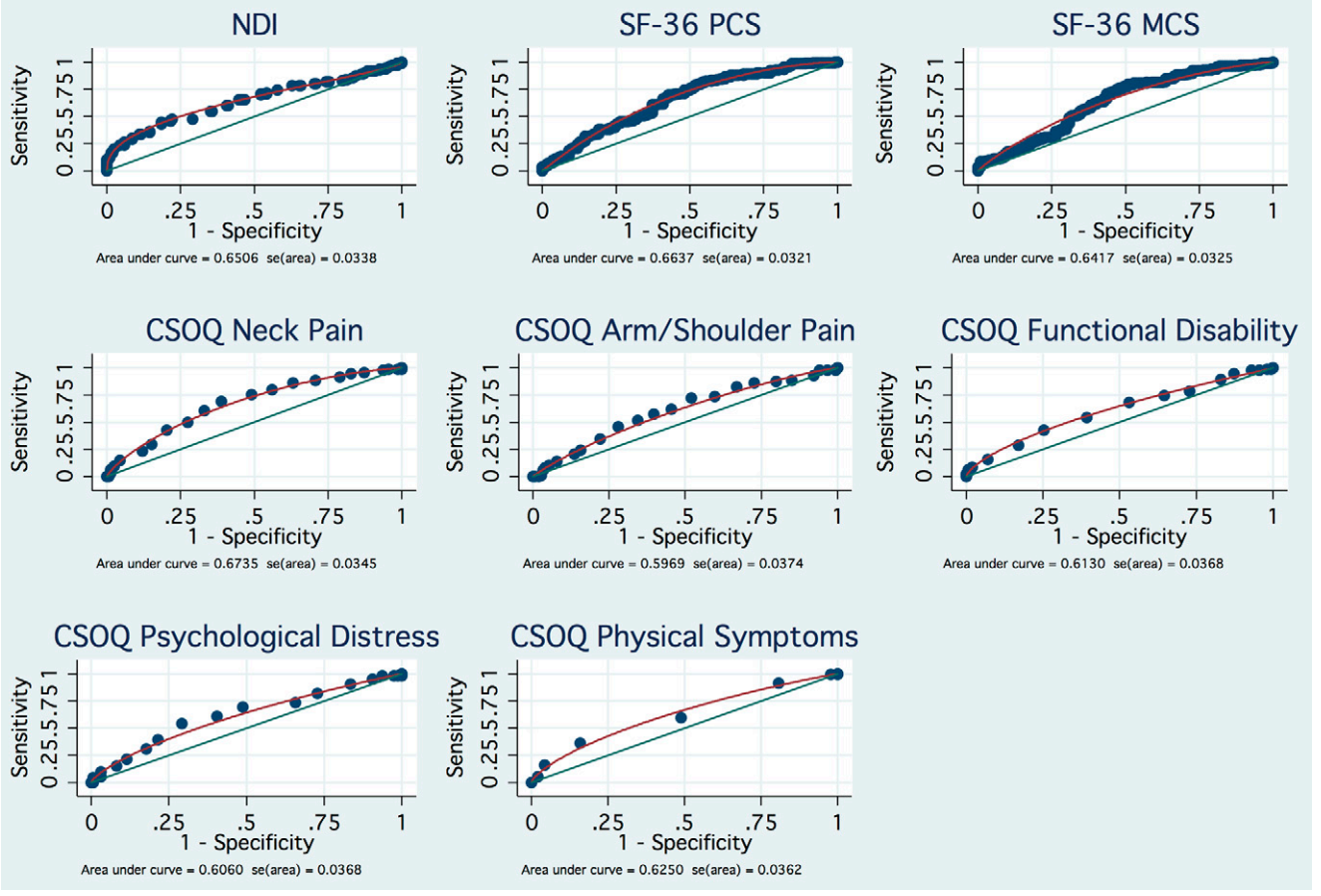

Fig. 1

Receiver operating characteristics curve matrix for the minimum clinically important difference values for the Neck Disability Index (NDI), Short Form-36 (SF-36) physical (PCS) and mental (MCS) health component scores, and Cervical Spine Outcomes Questionnaire (CSOQ) domain scores. For the NDI, each observed score was plotted with the associated sensitivity and specificity. The point nearest the upper left corner of the plot represents the score that maximizes sensitivity and specificity.

TABLE III Minimum Clinically Important Differences and Substantial Clinical Benefit Values for Patient-Reported Outcomes Instruments

Statistic	NDI* *(points)*	SF-36* *(points)*	
		Physical Component	Mental Component
Standard baseline	15.53	10.79	10.78
Standard error measure	6.95	2.85	2.85
Minimum clinically important difference	11.3 (0.641)	6.5 (0.659)	5.0 (0.639)
Substantial clinical benefit	13.0 (0.683)	7.8 (0.710)	6.6 (0.660)

*The number in parentheses is the area under the receiver operating characteristics curve.

Males and females demonstrated significant differences in the change in the CSOQ domain scores for neck pain (−0.24 versus −0.31, p = 0.022), but not for arm/shoulder pain (p = 0.908), functional disability (p = 0.586), psychologic distress (p = 0.092), or physical symptoms other than pain (p = 0.449) (Table IV). There were no significant correlations between the preoperative age and the change in CSOQ domain scores. There was a significant positive correlation between the duration of current symptoms and the change in the scores for neck pain (r = 0.127, p = 0.046) and arm/shoulder pain (r =

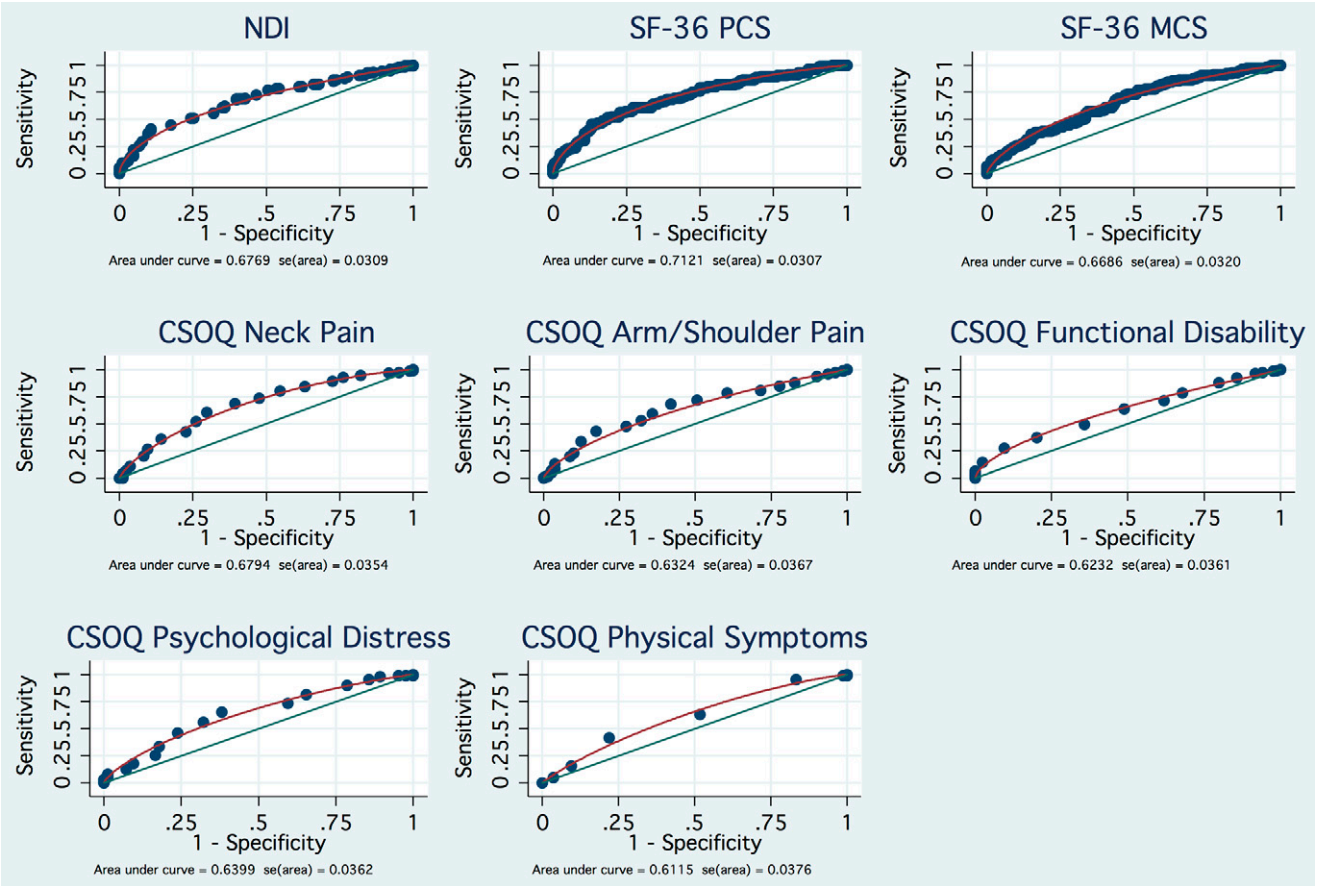

Fig. 2

Receiver operating characteristics curve matrix for the substantial clinical benefit values for the Neck Disability Index (NDI), Short Form-36 (SF-36) physical (PCS) and mental (MCS) health component scores, and Cervical Spine Outcomes Questionnaire (CSOQ) domain scores. For the NDI, each observed score was plotted with the associated sensitivity and specificity. The point nearest the upper left corner of the plot represents the score that maximizes sensitivity and specificity.

TABLE III (*continued*)

CSOQ* *(points)*				
Neck Pain	Arm/Shoulder Pain	Functional Disability	Psychologic Distress	Physical Symptoms Other Than Pain
0.23	0.20	0.21	0.21	0.30
0.07	0.06	0.07	0.07	0.09
0.20 (0.676)	0.24 (0.611)	0.13 (0.605)	0.17 (0.623)	0.20 (0.614)
0.30 (0.685)	0.30 (0.649)	0.20 (0.617)	0.22 (0.645)	0.20 (0.610)

TABLE IV Relationship Between Change in Cervical Spine Outcomes Questionnaire Domain Scores and Patient Sex

CSOQ Parameter	Mean CSOQ Domain Score* *(points)*		P Value
	Male	Female	
Neck pain	–0.24 ± 0.24	–0.31 ± 0.24	0.022
Arm/shoulder pain	–0.30 ± 0.29	–0.31 ± 0.25	0.908
Functional disability	–0.20 ± 0.22	–0.22 ± 0.21	0.586
Psychologic distress	–0.18 ± 0.19	–0.22 ± 0.23	0.092
Physical symptoms other than pain	–0.16 ± 0.26	–0.18 ± 0.27	0.449

*The values are given as the mean ± standard deviation.

0.166, p = 0.010), but not between the duration of current symptoms and the remaining domain scores (see Table E-1 in the Appendix).

Discussion

Outcome studies frequently rely on reports of the mean change on self-reported instruments measuring physical and mental health to denote the clinical benefit of a certain procedure. Without clear metrics against which to judge these changes to clinical practice, these outcomes studies have limitations. Minimum clinically important difference and substantial clinical benefit thresholds provide the metric by which researchers and clinicians can gauge mean changes in scores on self-reported instruments of health. The current study demonstrated that, among individuals who have undergone anterior cervical spine decompression and arthrodesis, the changes in the CSOQ domain scores representing a minimum clinically important difference and a substantial clinical benefit were 0.13 to 0.24 point and 0.20 to 0.30 point, respectively. In each case, these metrics exceed the standard error of measurement for these domain scores, indicating that the minimum clinically important difference and substantial clinical benefit values are greater than what could be accounted for by measurement error. Thus, these values represent a true change in the domain scores associated with clinical benefits. Women experienced a greater decrease in neck pain after surgery than did men, but we found no sex-related difference in the other CSOQ domain scores. Baseline preoperative age was not associated with the change in domain scores. An increase in the duration of current symptoms was associated with a smaller improvement in domain scores from baseline for neck pain and arm/shoulder pain, but not for the other domains.

There are three major methods for estimating the minimum clinically important difference and substantial clinical benefit: distribution-based, opinion-based, and anchor-based. With distribution-based methods, these clinical metrics are estimated on the basis of the distribution of scores in an untreated/unaffected population. Although these metrics are easy to calculate, limitations of this method include variability across studies and the absence of a clinical component[13]. Opinion-based methods are based on a consensus between patients and experts regarding minimum clinically important differences. Although expert opinion regarding changes in outcome measures may be useful, this method is rarely used. With anchor-based methods, an external patient-reported measure of change in status (e.g., current health state) is used to determine these metrics. We used the anchor-based method to determine the minimum clinically important difference and substantial clinical benefit with use of receiver operating characteristics curves.

Other investigators have estimated the minimum clinically important difference for the NDI[3,5]. The values that we calculated are higher than those reported by Cleland et al.[5]. It is important to note that those authors studied patients

undergoing nonoperative care for spinal abnormalities[5]. Differences among our study, the study by Cleland et al.[5], and the study by Carreon et al.[3] may be related to greater expectations for benefit on the part of patients undergoing surgical treatment, a longer follow-up time, or greater disability at baseline. Carreon et al. reported the minimum clinically important difference on the NDI in a population of patients who had undergone cervical spine surgery and had been followed for a minimum of one year after that surgery. Our findings indicate greater values for the minimum clinically important difference than those that they reported.

There are several limitations to our study. First, minimum clinically important difference and substantial clinical benefit values are specific to the populations in which they are developed. Thus, our findings are applicable to those undergoing anterior cervical spine decompression and arthrodesis for cervical degenerative disc disease but may not be applicable to patients undergoing nonoperative care or other types of surgical intervention. Second, we used baseline preoperative and six-month postoperative assessments to generate the minimum clinically important difference and substantial clinical benefit values. The thresholds reported in this study may not be appropriate for evaluation of shorter-term or longer-term change in domain scores on the CSOQ. Third, we did not examine threshold values to identify individuals who experienced a worsening of health status because of the limited number of patients who reported that finding when answering our anchor question.

Patients with a 0.20-point change in neck pain or a 0.24-change in arm/shoulder pain experienced a clinically important improvement in the self-reported outcome after surgery for cervical degenerative disc disease. Similarly, patients with a 0.30-point change in neck pain or arm/shoulder pain experienced a substantial clinical benefit after surgery.

Our study has provided metrics with which changes in the domain scores of the CSOQ can be compared to denote a minimum clinically important difference and a substantial clinical benefit. This instrument provides clinicians and researchers with an additional valid and reliable multidimensional outcomes measure with which to evaluate the progress of their patients after surgery. This information, coupled with previous reports of the psychometric stability of the CSOQ, should increase the clinical utility of this patient-reported outcomes instrument in cervical spine surgery.

Appendix

eA The Cervical Spine Outcomes Questionnaire (CSOQ) and a table showing the relationship between the change in the CSOQ domain scores and patient age and duration of current symptoms are available with the online version of this article at jbjs.org. ■

Richard L. Skolasky, ScD
Todd J. Albert, MD
Anica M. Maggard, BS
Lee H. Riley III, MD
Department of Orthopaedic Surgery,
Johns Hopkins Bayview Medical Center,
4940 Eastern Avenue, #A665, Baltimore, MD 21224.
E-mail address for R.L. Skolasky: ehenze1@jhmi.edu

References

1. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J. 2007;7:541-6.

2. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific Quality of Life Questionnaire. J Clin Epidemiol. 1994; 47:81-7.

3. 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;10:469-74.

4. Spratt KF. Patient-level minimal clinically important difference based on clinical judgment and minimally detectable measurement difference: a rationale for the SF-36 physical function scale in the SPORT intervertebral disc herniation cohort. Spine (Phila Pa 1976). 2009;34:1722-31.

5. Cleland JA, Fritz JM, Whitman JM, Palmer JA. The reliability and construct validity of the Neck Disability Index and patient specific functional scale in patients with cervical radiculopathy. Spine (Phila Pa 1976). 2006;31:598-602.

6. Young BA, Walker MJ, Strunce JB, Boyles RE, Whitman JM, Childs JD. Responsiveness of the Neck Disability Index in patients with mechanical neck disorders. Spine J. 2009;9:802-8.

7. Skolasky RL, Riley LH 3rd, Albert TJ. Psychometric properties of the Cervical Spine Outcomes Questionnaire and its relationship to standard assessment tools used in spine research. Spine J. 2007;7:174-9.

8. BenDebba M, Heller J, Ducker TB, Eisinger JM. Cervical spine outcomes questionnaire: its development and psychometric properties. Spine (Phila Pa 1976). 2002;27:2116-24.

9. Vernon H, Mior S. The Neck Disability Index: a study of reliability and validity. J Manipulative Physiol Ther. 1991;14:409-15.

10. McHorney CA, Ware JE Jr, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. Med Care. 1993;31:247-63.

11. McHorney CA, Ware JE Jr, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care. 1994;32:40-66.

12. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clin Chem. 1993;39:561-77.

13. Brozek JL, Guyatt GH, Schünemann HJ. 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. 2006;4:69.