J Neurosurg Spine 16:61–67, 2012

Determination of minimum clinically important difference in pain, disability, and quality of life after extension of fusion for adjacent-segment disease

Clinical article

Scott L. Parker, M.D.,[1,2] Stephen K. Mendenhall, B.S.,[1,2] David Shau, B.S.,[1,2] Owoicho Adogwa, M.P.H.,[1,2] Joseph S. Cheng, M.D., M.S.,[1,2] William N. Anderson, Ph.D.,[3] Clinton J. Devin, M.D.,[1,2] and Matthew J. McGirt, M.D.[1,2]

[1]*Departments of Neurosurgery and Orthopedic Surgery, Vanderbilt University Medical Center;* [2]*Vanderbilt University Spinal Column Surgical Quality and Outcomes Research Laboratory, Nashville, Tennessee; and* [3]*Independent Statistician, Lake Forest, California*

Object. Spinal surgical outcome studies rely on patient-reported outcome (PRO) measurements to assess treatment effect. A shortcoming of these questionnaires is that the extent of improvement in their numerical scores lack a direct clinical meaning. As a result, the concept of minimum clinical important difference (MCID) has been used to measure the critical threshold needed to achieve clinically relevant treatment effectiveness. As utilization of spinal fusion has increased over the past decade, so has the incidence of adjacent-segment degeneration following index lumbar fusion, which commonly requires revision laminectomy and extension of fusion. The MCID remains uninvestigated for any PROs in the setting of revision lumbar surgery for adjacent-segment disease (ASD).

Methods. In 50 consecutive patients undergoing revision surgery for ASD-associated back and leg pain, PRO measures of back and leg pain on a visual analog scale (BP-VAS and LP-VAS, respectively), Oswestry Disability Index (ODI), 12-Item Short Form Health Survey Physical and Mental Component Summaries (SF-12 PCS and MCS, respectively), and EuroQol-5D health survey (EQ-5D) were assessed preoperatively and 2 years postoperatively. The following 4 well-established anchor-based MCID calculation methods were used to calculate MCID: average change; minimum detectable change (MDC); change difference; and receiver operating characteristic curve (ROC) analysis for the following 2 separate anchors: health transition item (HTI) of the SF-36 and satisfaction index.

Results. All patients were available for 2-year PRO assessment. Two years after surgery, a statistically significant improvement was observed for all PROs (mean changes: BP-VAS score [4.80 ± 3.25], LP-VAS score [3.28 ± 3.25], ODI [10.24 ± 13.49], SF-12 PCS [8.69 ± 12.55] and MCS [8.49 ± 11.45] scores, and EQ-5D [0.38 ± 0.45]; all $p < 0.001$). The 4 MCID calculation methods generated a range of MCID values for each of the PROs (BP-VAS score, 2.3–6.5; LP-VAS score, 1.7–4.3; ODI, 6.8–16.9; SF-12 PCS, 6.1–12.6; SF-12 MCS, 2.4–10.8; and EQ-5D, 0.27–0.54). The area under the ROC curve was consistently greater for the HTI anchor than the satisfaction anchor, suggesting this as a more accurate anchor for MCID.

Conclusions. Adjacent-segment disease revision surgery–specific MCID is highly variable based on calculation technique. The MDC approach with HTI anchor appears to be most appropriate for calculation of MCID after revision lumbar fusion for ASD because it provided a threshold above the 95% CI of the unimproved cohort (greater than the measurement error), was closest to the mean change score reported by improved and satisfied patients, and was not significantly affected by choice of anchor. Based on this method, MCID following ASD revision lumbar surgery is 3.8 points for BP-VAS score, 2.4 points for LP-VAS score, 6.8 points for ODI, 8.8 points for SF-12 PCS, 9.3 points for SF-12 MCS, and 0.35 quality-adjusted life-years for EQ-5D. *(DOI: 10.3171/2011.8.SPINE1194)*

Key Words • **minimum clinically important difference** • **adjacent-segment disease** • **fusion**

Patient-reported outcome questionnaires have become the standard measure for treatment effectiveness following spinal surgery. The most commonly used PRO questionnaires include pain scales for back and leg pain (VAS),[8,10] ODI,[6,7,20] SF-36,[25] and EQ-5D.[1,14] A well-defined shortcoming of such questionnaires is that their numerical scores lack a direct, clinically significant

This article contains some figures that are displayed in color online but in black and white in the print edition.

Abbreviations used in this paper: ASD = adjacent-segment disease; AUC = area under the curve; BP-VAS = back pain–visual analog scale; EQ-5D = EuroQol-5D health survey; HTI = health transition item; LP-VAS = leg pain–VAS; MCID = minimum clinically important difference; MCS = Mental Component Summary; MDC = minimum detectable change; ODI = Oswestry Disability Index; PCS = Physical Component Summary; PRO = patient-reported outcome; QALY = quality-adjusted life-year; ROC = receiver operating characteristic; SF-12 = 12-Item Short Form Health Survey; SF-36 = 36-Item Short Form Health Survey; TLIF = transforaminal lumbar interbody fusion.

meaning.[17] As a result, the concept of MCID has been put forth as a measure for the critical threshold needed to achieve clinically meaningful treatment effectiveness. Using this measure, treatment effects reaching the MCID threshold value imply clinical significance and justification for implementation into clinical practice. In other words, MCID can be thought of as the smallest change in outcome measure that is important to patients.[22]

Anchor-based approaches are the most often used and most validated methods to calculate MCID values. An anchor-based approach compares the change in PRO score after surgery with another measurement (perceived improvement after surgery, satisfaction with surgery, and willingness to have surgery again given the experienced outcome). Unfortunately, multiple anchors have been used and several anchor-based MCID calculation methods have been described, resulting in substantial variability in MCID values.[24] Because of this variability, there has been no consensus on the best MCID calculation method; therefore, definitive MCID values have yet to be established for the aforementioned common PRO questionnaires used in spine surgery.

Minimum clinically important difference values are specific to the PRO metric being used, the spinal pathology, and the surgical procedure performed. Previous studies have attempted to determine the MCID of VAS,[11] ODI,[11] and SF-36[27,28] in mixed spine surgery populations of various etiologies and surgical procedures.[3,11] Copay et al.[4] assessed MCID for VAS, ODI, and EQ-5D in a patient population of mixed spine pathologies and surgeries and demonstrated a wide variability in MCID based on MCID calculation method. Our group[19] previously assessed MCID for VAS, ODI, and EQ-5D in patients undergoing TLIF for spondylolithesis-associated back and leg pain. To date, no studies have determined MCID values specifically for decompression and extension of fusion for symptomatic ASD. Therefore, we set out to determine the most appropriate ASD revision surgery–specific MCID values for VAS, ODI, SF-12, and EQ-5D in patients undergoing revision decompression and extension of fusion for symptomatic ASD.

Methods

Patient Selection

Fifty consecutive patients with a diagnosis of lumbar ASD who underwent revision fusion were included in this study. The primary inclusion criteria were prior instrumented lumbar fusion for degenerative lumbar stenosis, radiographic evidence of ASD, mechanical low-back and leg pain, an age of 18–70 years, and failure of at least 3 months of nonoperative therapy. Patients were excluded if they had an extraspinal cause of back and leg pain; evidence of trauma, infection, pseudarthrosis, or hardware failure; an active workers' compensation lawsuit; or were unwilling to participate with follow-up questions.

The diagnosis of ASD was made using the following criteria: low-back and leg pain localized to the adjacent level after prior fusion; and MR imaging or dynamic radiography evidence of any pathology immediately adjacent to the prior fusion segment including spinal stenosis, listhesis, or instability. Patients with same-level recurrent stenosis or pseudarthrosis at the level of the index surgery were not included.

Patient demographics, clinical presentation, indications for surgery, radiological studies, and operative variables were retrospectively reviewed for each case. Patient-assessed outcome measures were obtained via phone interview. All 4 surgeons participating in this study practiced similar postoperative treatment paradigms. In all cases, the surgeon encouraged discharge from the hospital beginning 48 hours after surgery, weaning of narcotics beginning 2–3 weeks after surgery, and return to work as soon as the patient felt capable.

Patient-Reported Outcome Measures

Institutional review board approval was obtained prior to contacting any patients for outcome assessment. Preoperative and 2-year outcomes after revision decompression and extension of fusion for symptomatic ASD were assessed by an independent investigator not involved with clinical care. Patient-reported outcome questionnaires included VAS scores for low-back pain (BP-VAS) and leg pain (LP-VAS),[8,16] ODI,[6,7] SF-12 PCS and MCS scores,[9,25] and EQ-5D[1,14,15] for quality of life. The EQ-5D was used for utility measurement as it has been validated and found to be responsive to low-back treatment.

Anchors

Two ad hoc anchors were used to demonstrate the anchor-based derivations of MCID. The first anchor used was derived from the HTI of the SF-36. This item asks patients to rate how their current health compares with their health prior to surgery. The choices provided include "Worse," "Unchanged," "Slightly Better," or "Markedly Better." Patients answering "slightly better" or "markedly better" were classified as responders, while those answering "unchanged" or "worse" were classified as nonresponders. The second anchor used was derived by asking patients whether they were satisfied with the results of their surgery. Patients answering "yes" were classified as responders, and those answering "no" were classified as nonresponders.

Anchor-Based Approaches

We chose 4 previously reported anchor-based approaches to assess MCID after revision decompression and extension of fusion for symptomatic ASD.[4,5] Each anchor-based approach calculates MCID based on the following unique definitions: 1) The "average change" approach defines MCID as the average change score seen in the cohort defined to be responders. 2) The MDC approach defines MCID as the smallest change that can be considered above the measurement error with a given level of confidence (often 95% confidence level); therefore, the MCID value is equal to the upper value of the 95% CI for average change score seen in the cohort defined to be nonresponders. 3) The "change difference" approach defines MCID as the difference of the average change score for responders and nonresponders. 4) The "ROC curve" approach defines MCID as the change value that provides the greatest sensitivity and/or specificity for a positive response. When used in the arena of MCID, sensitivity is defined as the proportion of patients who report improvement based on the external criterion and have a PRO score above the MCID threshold value. Conversely, specificity is defined as the proportion of patients who do not report improvement based on the external criterion and have a PRO score below the MCID threshold value.[5] The MCID value is defined as that which equally maximizes sensitivity and specificity. Additionally, the probability that scores will correctly discriminate between responders and nonresponders (accuracy) is depicted by the area under the ROC curve. This value ranges from 0.5 (discrimination is no better than pure chance) to 1.0 (all patients are able to be correctly discriminated). An AUC of 0.7–0.8 is considered adequate, while an AUC of 0.8–0.9 is considered excellent.[5]

Results

At our institution there were 2100 index lumbar fusions performed from 2004 through 2008. A total of 50 consecutive patients presenting with symptomatic ASD during this time were enrolled in the study. Based on our exclusion criteria, 100 patients undergoing revision fusion were excluded. Fifty-three patients undergoing revision, neural decompression, and fusion for same-level recurrent stenosis and 47 patients undergoing revision surgery for symptomatic lumbar pseudarthrosis were excluded. Twelve patients had an extraspinal cause of back pain, and 5 patients requiring revision lumbar surgery for postoperative infection were also excluded.

All patients in this series underwent a posterior fusion procedure with extension of the prior fusion construct. No patients within the series were excluded or lost to follow-up, and all patients had a minimum follow-up of 2 years. Overall, the mean age ± SD was 58.9 ± 10.9 years (29 women and 21 men). Seven patients (14%) had diabetes and 3 (6%) were either current or previous smokers. The mean ± SD body mass index was 29.5 ± 4.8.

All patients presented with back and leg pain as well as radiographic evidence of ASD. The indication for surgery was spondylosis/spinal stenosis in 36 patients (72%), spondylolisthesis in 13 patients (26%), and spinal instability in 1 patient (2%). The mean ± SD time between index lumbar fusion and revision surgery for ASD was 5.8 ± 7.0 years. The number of levels fused during revision surgery was 3.1 ± 0.8. Fifteen patients (30%) underwent 2-level fusion, 17 (34%) underwent 3-level fusion, and 18 (36%) underwent 4-level fusion. Twenty-eight patients (56%) had spinal fusion at L2–3, 38 (76%) at L3–4, 50 (100%) at L4–5, and 37 (74%) at L5–S1. All patients had a laminectomy performed at the adjacent levels, and 25 (50%) had an interbody graft placed. The ASD was located above the original level in 37 patients (74%), below in 9 patients (18%), and on both sides in 4 patients (8%).

At the time of presentation, the mean BP-VAS and LP-VAS were 8.7 ± 1.9 and 6.3 ± 3.9, respectively. Preoperative ODI and EQ-5D were 28.7 ± 9.6 and 0.29 ± 0.28, respectively. The preoperative SF-12 PCS and MCS scores were 26.9 ± 8.9 and 44.7 ± 12.9, respectively.

At 2 years postoperatively, each of the outcome measures assessed had significantly improved. Two years after surgery, the mean improvement in BP-VAS and LP-VAS was 4.80 ± 3.25 and 3.28 ± 3.25, respectively (both $p < 0.001$). The mean improvement in ODI, SF-12 PCS, and SF-12 MCS scores were 10.24 ± 13.49, 8.69 ± 12.55, and 8.49 ± 11.45, respectively (all $p < 0.001$; Fig. 1). The mean change in EQ-5D US–scaled index was 0.38 ± 0.45 QALYs ($p < 0.001$).

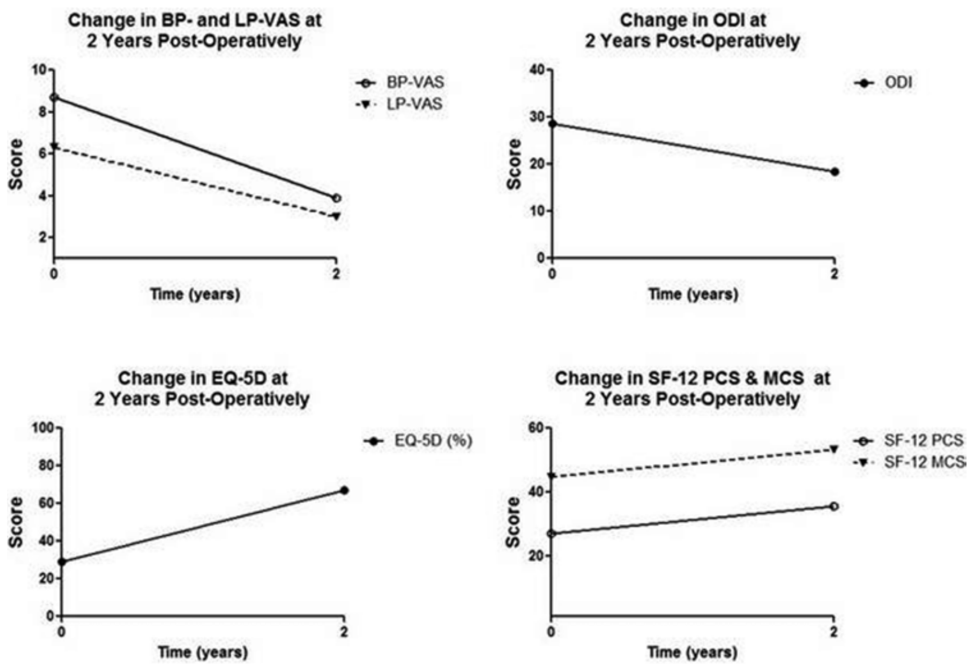

Fig. 1. Baseline and 2-year scores for patients undergoing revision decompression and extension of fusion for symptomatic ASD. Each PRO questionnaire revealed significant improvement at 2 years postoperatively. The mean change score ± SD for BP-VAS, LP-VAS, ODI, SF-12 PCS, SF-12 MCS, and EQ-5D was 4.80 ± 3.25, 3.28 ± 3.25, 10.24 ± 13.49, 8.69 ± 12.55, 8.49 ± 11.45, and 0.38 ± 0.45, respectively (all $p < 0.001$).

For the HTI anchor, 33 patients (66%) were classified as responders and 17 (34%) as nonresponders. For the satisfaction with surgery anchor, 24 patients (48%) answered "yes" and 26 (52%) answered "no."

Visual Analog Scale for Low-Back Pain

Based on calculation method, the MCID threshold ranged from 3.5 to 6.0 for the HTI anchor and 2.3 to 6.0 for the satisfaction with surgery anchor. For both anchors, the smallest threshold was derived from the "change difference" approach and the largest from the "average change" approach. The area under the ROC curve was greater for the HTI anchor (0.79 vs 0.69; Table 1).

Visual Analog Scale for Leg Pain

The MCID threshold ranged from 2.4 to 4.3 for the HTI anchor and 1.7 to 4.2 for the satisfaction with surgery anchor. For the HTI anchor, the smallest threshold was derived from the MDC approach, while the "change difference" approach yielded the smallest threshold for the satisfaction with surgery anchor. For both anchors, the largest threshold was derived from the "average change" approach. The area under the ROC curve was greater for the HTI anchor (0.78 vs 0.65; Table 1).

Oswestry Disability Index

The MCID threshold ranged from 6.8 to 14.9 for the HTI anchor and 9.9 to 16.7 for the satisfaction with surgery anchor. For both anchors, the smallest threshold was derived from the MDC approach and the largest from the "average change" approach. The area under the ROC curve was greater for the HTI anchor (0.80 vs 0.79; Table 1).

12-Item Short Form Health Survey PCS

The MCID threshold ranged from 6.2 to 11.7 for the HTI anchor and 7.5 to 12.6 for the satisfaction with surgery anchor. For the HTI anchor, the smallest threshold was derived from the "ROC curve" approach, while the "change difference" approach yielded the smallest threshold for the satisfaction with surgery anchor. For both anchors, the largest threshold was derived from the "average change" approach. The area under the ROC curve was greater for the HTI anchor (0.71 vs 0.69; Table 1).

12-Item Short Form Health Survey MCS

The MCID threshold ranged from 6.3 to 10.6 for the HTI anchor and 4.4 to 10.8 for the satisfaction with surgery anchor. For both anchors, the smallest threshold was derived from the "change difference" approach and the largest from the "average change" approach. The area under the ROC curve was greater for the HTI anchor (0.64 vs 0.61; Table 1).

EuroQuol-5D Health Survey

The MCID threshold ranged from 0.29 to 0.48 for the HTI anchor and 0.29 to 0.53 for the satisfaction with surgery anchor. For both anchors, the smallest threshold was derived from the "change difference" approach and the largest from the "average change" approach. The area under the ROC curve was greater for the HTI anchor (0.75 vs 0.66; Table 1).

Comparison of Anchor and MCID Calculation

Similar to variations in MCID values between calculation methods, the MCID threshold value varied to a lesser degree based on the anchor used (Table 1). The area under the ROC curve was greater for the HTI versus satisfaction with surgery anchor for all PRO measures evaluated (Fig. 2). When comparing the average AUC from the 6 PRO-specific ROC calculations, the HTI (AUC 0.75) versus satisfaction anchor (AUC 0.68) was more accurate, suggesting HTI to be a more valid anchor for ASD revision surgery-specific MCID. For each of the PROs assessed, the "average change" approach consistently produced the largest MCID value, while the "change difference" approach most often produced the smallest MCID value.

The MDC approach for calculating MCID appears to be the most appropriate MCID value in this patient cohort because it provided a threshold above the 95% CI of the unimproved cohort (greater than the measurement error), was closest to the mean change score reported by improved patients, and was least affected by the choice of anchor. Based on the MDC method with the HTI anchor, MCID following revision decompression and extension of fusion for symptomatic ASD are 3.8 points for BP-VAS, 2.4 points for LP-VAS, 6.8 points for ODI, 8.8 points for SF-12 PCS,

TABLE 1: Assessment of outcome factors*

Anchor-Based Approach	Calculated MCID											
	BP-VAS		LP-VAS		ODI		SF-12 PCS		SF-12 MCS		EQ-5D	
	HTI	Satisfied	HTI	Satisfied	HTI	Satisfied	HTI	Satisfied	HTI	Satisfied	HTI	Satisfied
average change	6.0	6.0	4.3	4.2	14.9	16.7	11.7	12.6	10.6	10.8	0.48	0.53
MDC (95% CI)	3.8	4.9	2.4	3.5	6.8	9.9	8.8	10.3	9.3	10.8	0.35	0.45
change difference	3.5	2.3	2.9	1.7	13.8	12.4	8.8	7.5	6.3	4.4	0.29	0.29
ROC curve	5.0	6.0	3.0	4.0	11.0	12.0	6.2	10.1	7.3	10.0	0.40	0.49
AUC	0.79	0.69	0.78	0.65	0.80	0.79	0.71	0.69	0.64	0.61	0.75	0.66

* The 4 MCID calculations generated a range of values for each of the PROs assessed. The HTI of the SF-36 was used as one anchor and compared with a second anchor based on satisfaction with results of surgery. The area under the ROC curve was consistently greater for the HTI anchor rather than patient satisfaction anchor, suggesting this as a more accurate anchor for MCID. The MDC approach may be the most appropriate MCID calculation based on the fact that it generated a threshold of improvement that was statistically greater than chance error from unimproved patients, was most consistent with the difference in change scores between responders and nonresponders, and was not significantly affected by choice of anchor.

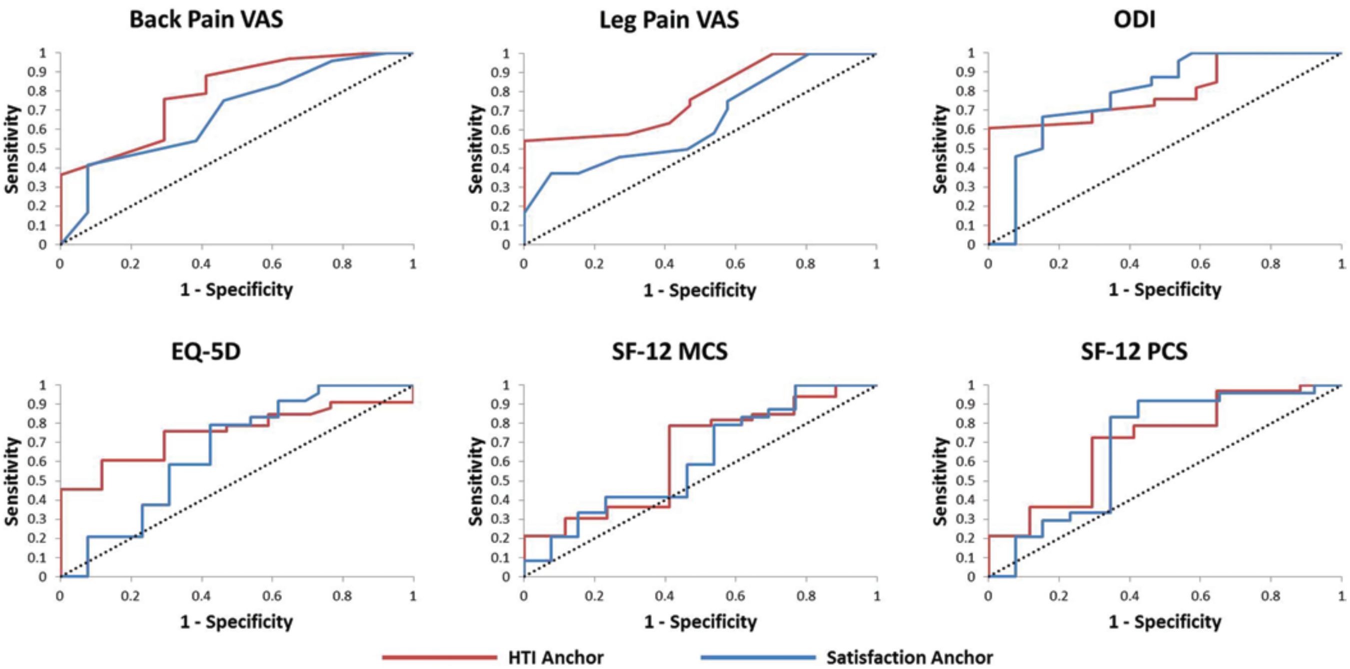

Fig. 2. Graph depicting the area under the ROC curve for the HTI and satisfaction anchors. The area under the ROC curve was greater for the HTI versus satisfaction with surgery anchor for all PRO measures evaluated.

9.3 points for SF-12 MCS, and 0.35 QALYs for EQ-5D. In this series, the mean patient change score surpassed the MCID threshold for each of the PRO measures assessed except SF-12 PCS and MCS (Fig. 3).

Discussion

In this study we set out to determine whether MCID thresholds varied as a function of MCID calculation method in a homogeneous patient population and to define an optimal MCID value for 6 commonly used PRO questionnaires in patients undergoing revision decompression and extension of fusion for symptomatic ASD. Each of the 4 anchor-based approaches used to calculate an MCID value resulted in a significant range of threshold values. The MCID threshold values ranged from 2.3 to 6.5 for BP-VAS, 1.7 to 4.3 for LP-VAS, 6.8 to 16.9 for ODI, 6.1 to 12.6 for SF-12 PCS, 2.4 to 10.8 for SF-12 MCS, and 0.27 to 0.54 for EQ-5D. The largest threshold value consistently resulted from the "average change" approach, and the smallest threshold most often came from the "change difference" approach. All 4 calculation methods assessed here are currently used in the literature. These results highlight the need to standardize the most appropriate calculation method for MCID.

Interestingly, only 48% of patients labeled themselves as being satisfied with the results of their surgery. The fact that approximately half of the patients were not satisfied with their surgery in light of significant improvements in PRO measures highlights the fact that current outcome measurements may be a poor proxy for patient satisfaction. Scientifically quantifiable improvements in outcome questionnaires or radiological fusion rates do not necessarily correlate with a patient's intrinsic view of a meaningful improvement. Additionally, the definition regarding what is a meaningful extent of improvement may vary from patient to patient. The multifactorial reason for a patient to be intrinsically satisfied following a surgical procedure remains unclear and underscores the importance of performing further similar analyses on other patient populations.

We also compared 2 anchors (general health assessment in the form of the HTI of the SF-36 and satisfaction with surgery). Traditionally, general health assessments have been used as anchors and have been shown to be sensitive to change.[11] It is also widely used due to the convenience associated with its use, as it is part of the commonly used SF-36. In the current study, there was moderate variability in the MCID threshold between the 2 anchors, with the HTI anchor consistently displaying a smaller MCID value. The HTI anchor was associated with a greater AUC for all 6 PRO metrics analyzed, suggesting a superior sensitivity and specificity compared with the satisfaction with surgery anchor. As a result, the HTI anchor appears to be the more valid anchor in this study.

To date, no consensus has been reached as to the superior method with which to calculate MCID. Calculation methods that have been previously supported in the literature include standard error of measurement,[13,27,28] one-half standard deviation,[2,18,26] effect size,[12,17,21,23] MDC,[4,19] and ROC curve.[3] The MCID study performed by Copay et al.[4] also systematically compared calculation methods and their resulting values within the same patient population. Similar to the results of the current study, a wide variability in MCID based on MCID calculation method was demonstrated. Based on the assertion that a truly sound MCID value should be at least greater than the measurement error and correspond to the patient perception of importance of change, both studies showed MDC to be the superior calculation method for determining an MCID threshold in patients undergoing various lumbar spine surgeries.

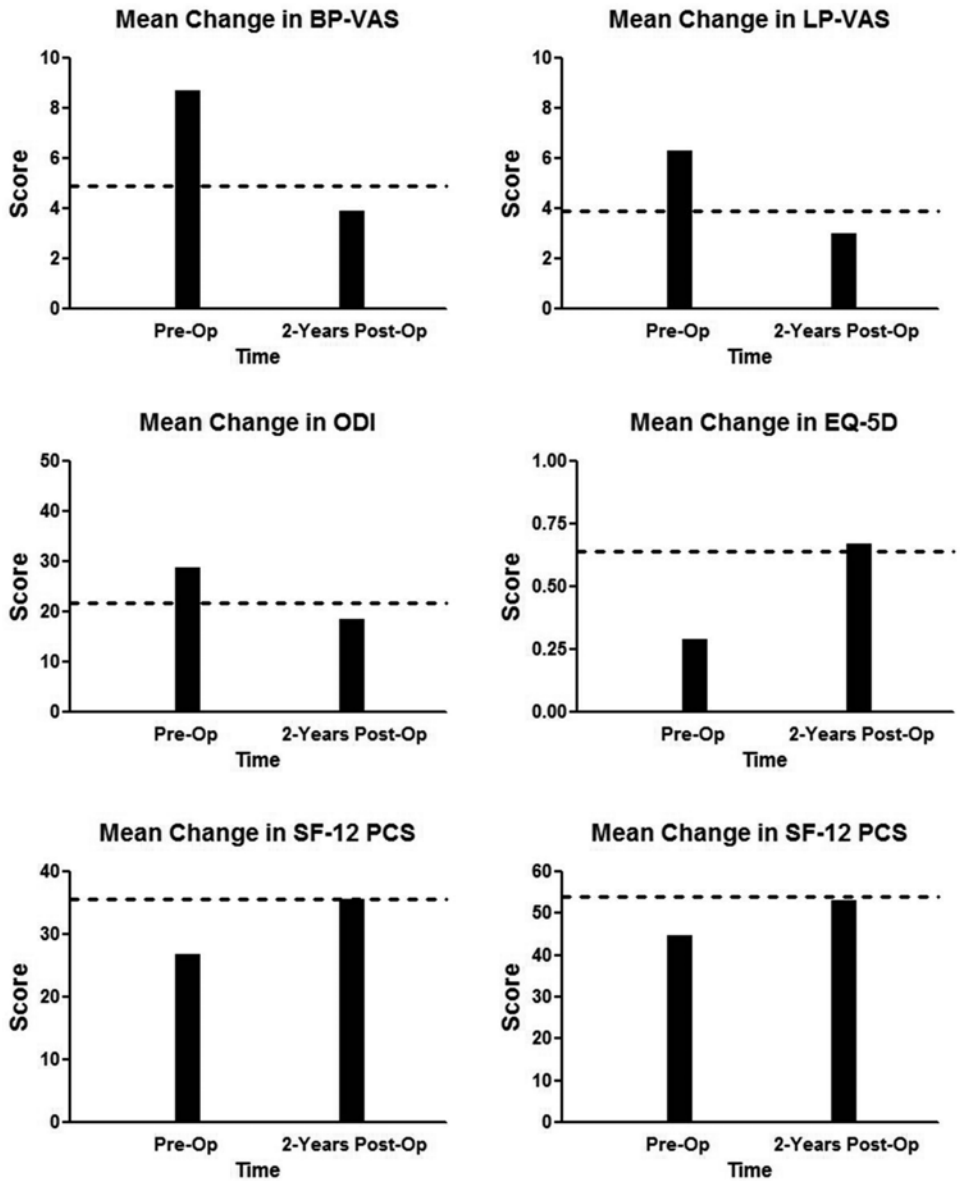

Fig. 3. Bar graph depicting the mean baseline and 2-year postoperative scores for patients undergoing revision decompression and extension of fusion for symptomatic ASD. The *dashed line* in each graph represents the threshold change score needed to attain MCID. Based on the MDC approach with the HTI anchor, the MCID thresholds are 3.8 points for BP-VAS, 2.4 points for LP-VAS, 6.8 points for ODI, 8.8 points for SF-12 PCS, 9.3 points for SF-12 MCS, and 0.35 QALYs for EQ-5D. In this series, the mean patient change score surpassed the MCID threshold for all PROs assessed except for SF-12 PCS and MCS.

Our group has previously defined MCID following TLIF in patients with low-grade degenerative lumbar spondylolisthesis.[19] Similar to the current study, the MDC method with the HTI anchor was determined to be the most appropriate calculation technique. The MCID following TLIF was 2.1 points for BP-VAS, 2.8 points for LP-VAS, 14.9 points for ODI, and 0.46 QALYs for EQ-5D. It is imperative to recognize the importance and utility of assessing MCID for a homogeneous patient population undergoing a specific surgical procedure. Each of these patient populations will be associated with unique pain perception as well as emotional and physical stressors. These differences will be elicited in the PRO measures. The revision fusion patient population described in the current study is different from those presenting with de novo back pain seeking primary fusion; therefore, the slight differences in MCID observed here should not be unexpected.

Conclusions

Adjacent-segment disease revision surgery–specific MCID is highly variable based on calculation technique. The MDC approach with HTI anchor appears to be most appropriate for calculation of MCID after revision lumbar fusion for ASD because it provided a threshold above the 95% CI of the unimproved cohort (greater than the measurement error), was closest to the mean change score reported by improved and satisfied patients, and was not significantly affected by choice of anchor. Based on this method, MCID following ASD revision lumbar surgery is 3.8 points for BP-VAS, 2.4 points for LP-VAS, 6.8 points for ODI, 8.8 points for SF-12 PCS, 9.3 points for SF-12 MCS, and 0.35 QALYs for EQ-5D.

Disclosure

The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

Author contributions to the study and manuscript preparation include the following. Conception and design: McGirt. Acquisition of data: Mendenhall, Shau, Adogwa. Analysis and interpretation of data: Parker. Drafting the article: Parker. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: McGirt. Statistical analysis: Anderson.

References

1. Badia X, Diaz-Prieto A, Gorriz MT, Herdman M, Torrado H, Farrero E, et al: Using the EuroQol-5D to measure changes in quality of life 12 months after discharge from an intensive care unit. **Intensive Care Med 27:**1901–1907, 2001
2. Beaton DE: Simple as possible? Or too simple? Possible limits to the universality of the one half standard deviation. **Med Care 41:**593–596, 2003
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 10:**469–474, 2010
4. 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 8:**968–974, 2008
5. 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 7:**541–546, 2007
6. Fairbank JC, Couper J, Davies JB, O'Brien JP: The Oswestry low back pain disability questionnaire. **Physiotherapy 66:**271–273, 1980
7. Fairbank JC, Pynsent PB: The Oswestry Disability Index. **Spine (Phila Pa 1976) 25:**2940–2952, 2000
8. Gallagher EJ, Liebman M, Bijur PE: Prospective validation of clinically important changes in pain severity measured on a visual analog scale. **Ann Emerg Med 38:**633–638, 2001
9. Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, et al: Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. International Quality of Life Assessment. **J Clin Epidemiol 51:**1171–1178, 1998
10. Grönblad M, Hupli M, Wennerstrand P, Järvinen E, Lukinmaa A, Kouri JP, et al: Intercorrelation and test-retest reliability of the Pain Disability Index (PDI) and the Oswestry Disability Questionnaire (ODQ) and their correlation with pain intensity in low back pain patients. **Clin J Pain 9:**189–195, 1993
11. Hägg O, Fritzell P, Nordwall A: The clinical importance of changes in outcome scores after treatment for chronic low back pain. **Eur Spine J 12:**12–20, 2003
12. Hays RD, Woolley JM: The concept of clinically meaningful difference in health-related quality-of-life research. How meaningful is it? **Pharmacoeconomics 18:**419–423, 2000
13. Jaeschke R, Singer J, Guyatt GH: Measurement of health status. Ascertaining the minimal clinically important difference. **Control Clin Trials 10:**407–415, 1989
14. Jansson KA, Németh G, Granath F, Jönsson B, Blomqvist P: Health-related quality of life (EQ-5D) before and one year after surgery for lumbar spinal stenosis. **J Bone Joint Surg Br 91:**210–216, 2009
15. Johnson JA, Coons SJ, Ergo A, Szava-Kovats G: Valuation of EuroQOL (EQ-5D) health states in an adult US sample. **Pharmacoeconomics 13:**421–433, 1998
16. Kelly AM: The minimum clinically significant difference in visual analogue scale pain score does not differ with severity of pain. **Emerg Med J 18:**205–207, 2001
17. Kulkarni AV: Distribution-based and anchor-based approaches provided different interpretability estimates for the Hydrocephalus Outcome Questionnaire. **J Clin Epidemiol 59:**176–184, 2006
18. Norman GR, Sloan JA, Wyrwich KW: Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. **Med Care 41:**582–592, 2003
19. Parker SL, Adogwa O, Paul AR, Anderson WN, Aaronson O, Cheng JS, et al: Utility of minimum clinically important difference in assessing pain, disability, and health state after transforaminal lumbar interbody fusion for degenerative lumbar spondylolithesis. Clinical article. **J Neurosurg Spine 14:** 598–604, 2011
20. Roland M, Fairbank J: The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. **Spine (Phila Pa 1976) 25:**3115–3124, 2000
21. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D: Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. **Pharmacoeconomics 15:** 141–155, 1999
22. Stratford PW, Binkley JM, Riddle DL, Guyatt GH: Sensitivity to change of the Roland-Morris Back Pain Questionnaire: part 1. **Phys Ther 78:**1186–1196, 1998
23. Taylor SJ, Taylor AE, Foy MA, Fogg AJ: Responsiveness of common outcome measures for patients with low back pain. **Spine (Phila Pa 1976) 24:**1805–1812, 1999
24. van der Roer N, Ostelo RW, Bekkering GE, van Tulder MW, de Vet HC: 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, 2006
25. Ware JE Jr: SF-36 health survey update. **Spine (Phila Pa 1976) 25:**3130–3139, 2000
26. Wright JG: Interpreting health-related quality of life scores: the simple rule of seven may not be so simple. **Med Care 41:**597–598, 2003
27. Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD: Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. **Med Care 37:**469–478, 1999
28. Wyrwich KW, Tierney WM, Wolinsky FD: Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. **J Clin Epidemiol 52:**861–873, 1999

Manuscript submitted February 5, 2011.
Accepted August 22, 2011.
Please include this information when citing this paper: published online September 30, 2011; DOI: 10.3171/2011.8.SPINE1194.
Address correspondence to: Matthew J. McGirt, M.D., 4347 Village at Vanderbilt, Nashville, Tennessee 37232-8618. email: matt.mcgirt@vanderbilt.edu.