ELSEVIER

The SPINE JOURNAL

The Spine Journal 000 (2023) 1–11

Clinical Study

Determining minimal clinically important difference estimates following surgery for degenerative conditions of the lumbar spine: analysis of the Canadian Spine Outcomes and Research Network (CSORN) registry

J. Denise Power, PhD[#,a], Anthony V. Perruccio, PhD[#,a,b,c], Mayilee Canizares, PhD[a], Greg McIntosh, MSc[d], Edward Abraham, MD[e], Najmedden Attabib, MD[f], Christopher S. Bailey, MD, MSc[g], Raphaële Charest-Morin, MD[h], Nicholas Dea, MD[h], Joel Finkelstein, MD, MSc[i], Charles Fisher, MD, MHSc[h], R. Andrew Glennie, MD[j], Hamilton Hall, MD[k], Michael G. Johnson, MD[l], Adrienne M. Kelly, MD[m], Stephen Kingwell, MD[n], Neil Manson, MD[e], Andrew Nataraj, MD[o], Jérôme Paquet, MD, PhD[p], Supriya Singh, MD[g], Alex Soroceanu, MD[q], Kenneth C Thomas, MD[q], Michael H. Weber, MD, PhD[r], Y. Raja Rampersaud, MD[a,b,*]

[a] *Schroeder Arthritis Institute, Krembil Research Institute, University Health Network, 399 Bathurst St, Toronto, Ontario, M5T 2S8, Canada*

[b] *Orthopaedics, Department of Surgery, University of Toronto, 149 College St, Toronto, Ontario, M5T 1P5, Canada*

[c] *Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, 155 College St 4th floor, Toronto, Ontario, M5T 3M6, Canada*

[d] *Canadian Spine Outcomes and Research Network, 10 Armstrong Cresent, Markdale, Ontario N0C 1H0, Canada*

[e] *Canada East Spine Center and Horizon Health Network, 400 University Avenue, 3C South, Saint John, New Brunswick, E2L 4L4, Canada*

[f] *Canada East Spine Centre, Division of Neurosurgery, Zone 2, Horizon Health Network, 400 University Avenue, 3B North, Saint John, New Brunswick E2L 4L4, Canada*

[g] *London Health Science Centre Combined Orthopaedic and Neurosurgery Spine Program, Schulich School of Medicine, Western University, 1151 Richmond Street, London, Ontario N6A 5C1, Canada*

[h] *Combined Neurosurgery and Orthopaedic Spine Program, University of British Columbia, 5959 Student Union Blvd, Vancouver, British Columbia V6T 1K2, Canada*

[i] *Spine Program, Sunnybrook Health Sciences Centre, 2075 Bayview Ave, Toronto, Ontario M4N 3M5, Canada*

[j] *Department of Surgery, Division of Orthopedics, Dalhousie University, 6230 Coburg Rd, Halifax, Nova Scotia B3H 4J5, Canada*

[k] *Department of Surgery, University of Toronto, 149 College Street, Room 508-A, Toronto, Ontario M5T 1P5, Canada*

FDA device/drug status: Not applicable.

Author disclosures: ***JDP:*** Nothing to disclose; ***AVP:*** Nothing to disclose; ***MC:*** Nothing to disclose; ***GM:*** Nothing to disclose; ***EA:*** Consulting: Medtronic Canada (B), Grant: Medtronic Canada (F); ***NA:*** Research Support (Investigator Salary, Staff/Materials) Rick Hansen Institute (none); ***CSB:*** Nothing to disclose; ***RCM:*** Nothing to disclose; ***ND:*** Stock Ownership: Medtronic (A), Consulting: Striker (D), Consulting: Baxter (B), Speaking and/or Teaching Arrangements: Medtronic (B); ***JF:*** Fellowship Support: Stryker Spine Canada (E Paid directly to institution/employer), Zimmer Biomet (C Paid directly to institution/employer); ***CGF:*** Royalties: Medtronic Canada (none), Consulting, Medtronic Canada, Nuvasive (none), Grants: OREF (none), Fellowship Support: Medtronic Canada, AO Spine (none); ***RAG:*** Grants: Medtronic Canada (E Paid directly to institution/employer); ***HH:*** Nothing to disclose; ***MGJ:*** Research Support (Investigator Salary, Staff/Materials): Stryker (F Paid directly to institution/employer); ***AMK:*** Nothing to disclose; ***SK:*** Nothing to disclose; ***NM:*** Consulting: Medtronic Canada (B); Grants: Medtronic Canada (F Paid directly to institution/employer); ***AN:*** Nothing to disclose; ***JP:*** Nothing to disclose; ***SS:*** Nothing to disclose; ***AS:*** Nothing to disclose; ***KCT:*** Nothing to disclose; ***MHW:*** Nothing to disclose; ***YRR:*** Royalties: Medtronic Canada (D), Consulting: Medtronic Canada (B).

*Corresponding author. Toronto Western Hospital, University Health Network, 399 Bathurst St, East Wing, 1st Floor, Room 1-441, Toronto, Ontario, Canada, M5T 2S8.

E-mail address: raja.rampersaud@uhn.ca (Y.R. Rampersaud).

[#] Joint first authorship.

https://doi.org/10.1016/j.spinee.2023.05.001

[l] Winnipeg Spine Program Health Sciences Centre, University of Manitoba, 820 Sherbrook Street, Winnipeg, Manitoba R3A 1R9, Canada
[m] Sault Area Hospital, Northern Ontario School of Medicine, 750 Great Northern Rd, Sault Ste. Marie, Ontario P6B 0A8, Canada
[n] Division of Orthopaedic Surgery, University of Ottawa, 145 Jean-Jacques-Lussier Private, Ottawa, Ontario K1N 6N5, Canada
[o] Division of Neurosurgery, University of Alberta, 8440 112 St NW, Edmonton, Alberta T6G 2B7, Canada
[p] Hôpital de L'Enfant-Jésus, Laval University, 401 18E Rue, Quebec City, Quebec G1J 1Z4, Canada
[q] University of Calgary Spine Program, University of Calgary, 2500 University Dr, NW, Calgary, Alberta T2N 1N4, Canada
[r] Division of Orthopaedics, McGill University Health Centre, 3650 Rue Saint-Urbain Bureau K 124, Montreal, Quebec H2X 2P4, Canada

Received 6 October 2022; revised 24 April 2023; accepted 2 May 2023

ABSTRACT

BACKGROUND CONTEXT: There is significant variability in minimal clinically important difference (MCID) criteria for lumbar spine surgery that suggests population and primary pathology specific thresholds may be required to help determine surgical success when using patient reported outcome measures (PROMs).

PURPOSE: The purpose of this study was to estimate MCID thresholds for 3 commonly used PROMs after surgical intervention for each of 4 common lumbar spine pathologies.

STUDY DESIGN/SETTING: Observational longitudinal study of patients from the Canadian Spine Outcomes and Research Network (CSORN) national registry.

PATIENT SAMPLE: Patients undergoing surgery from 2015 to 2018 for lumbar spinal stenosis (LSS; n = 856), degenerative spondylolisthesis (DS; n = 591), disc herniation (DH; n = 520) or degenerative disc disease (DDD n = 185) were included.

OUTCOME MEASURES: PROMs were collected presurgery and 1-year postsurgery: the Oswestry Disability Index (ODI), and back and leg Numeric Pain Rating Scales (NPRS). At 1-year, patients reported whether they were 'Much better'/'Better'/'Same'/'Worse'/'Much worse' compared to before their surgery. Responses to this item were used as the anchor in analyses to determine surgical MCIDs for benefit ('Much better'/'Better') and substantial benefit ('Much better').

METHODS: MCIDs for absolute and percentage change for each of the 3 PROMs were estimated using a receiving operating curve (ROC) approach, with maximization of Youden's index as primary criterion. Area under the curve (AUC) estimates, sensitivity, specificity and correct classification rates were determined. All analyses were conducted separately by pathology group.

RESULTS: MCIDs for ODI change ranged from −10.0 (DDD) to −16.9 (DH) for benefit, and −13.8 (LSS) to −22.0 (DS,DH) for substantial benefit. MCID for back and leg NPRS change were −2 to −3 for each group for benefit and −4.0 for substantial benefit for all groups on back NPRS. MCID estimates for percentage change varied by PROM and pathology group, ranging from −11.1% (ODI for DDD) to −50.0% (leg NPRS for DH) for benefit and from −40.0% (ODI for DDD) to −66.6% (leg NPRS for DH) for substantial benefit. Correct classification rates for all MCID thresholds ranged from 71% to 89% and were relatively lower for absolute vs. percent change for those with high or low presurgical scores.

CONCLUSIONS: Our findings suggest that the use of generic MCID thresholds across pathologies in lumbar spine surgery is not recommended. For patients with relatively low or high presurgery PROM scores, MCIDs based on percentage change, rather than absolute change, appear generally preferable. These findings have applicability in clinical and research settings, and are important for future surgical prognostic work.

Keywords: Disability; Lumbar spine; MCID; Minimal clinically important difference; Pain; Surgery

Introduction

Patient-reported outcome measures (PROMs) are increasingly employed to assess clinical outcomes after lumbar spine surgery, and to inform clinical decision-making and health policy. PROMs help facilitate an understanding of whether patients perceive important improvements in spine-related symptoms such as pain and function. However, during longitudinal study some difficulty remains with interpreting the meaning of a change in numerical score on a given PROM, as the correlations between change scores and measures of patient-reported

postsurgical satisfaction or outcome are highly variable [1,2]. The concept of a minimal clinically importance difference (MCID) has been increasingly employed to quantify the minimal change in a given PROM that is considered clinically relevant [3].

There have been a number of lumbar spine-specific MCIDs developed for commonly used PROMs such as the Oswestry Disability Index (ODI) and back and leg numeric pain rating scales (NPRS) [4−11]. Unfortunately, there is significant variability among these MCID estimates. Studies used to generate these estimates have varied in terms of the lumbar spine pathologies considered and the methodologies employed. Some have used distribution-type methods that are very sample-specific, depending highly on the statistical characteristics of the study sample, rather than considering the relationship between the degree of change experienced and any measures of patient perceived outcome or satisfaction [12−16]. Whereas others have only considered postsurgical thresholds without accounting for presurgical score. Further, although there is literature to support that percentage change estimates may be preferable to estimates of absolute change, particularly for patients with low or high presurgery scores [17,18], MCID estimates for percentage change have been determined in only a very limited number of lumbar spine surgery populations [6,7].

The purpose of this study was to estimate thresholds of important change after surgical intervention for each of four common lumbar spine pathologies: lumbar spinal stenosis (LSS), degenerative spondylolisthesis (DS), disc herniation (DH) or degenerative disc disease (DDD). We used an anchor-based approach that explicitly considered patient-perceived outcome. We estimated thresholds for two levels of surgical improvement for both absolute and percentage change in scores on three frequently used PROMs. To the best of our knowledge, this is the first study to determine MCID estimates using data from a Canadian surgical lumbar spine sample. Given the high and increasing volume and cost of lumbar spine procedures, the development of such thresholds is an important step in understanding surgical success and its predictors, with the goal of improving patient outcomes.

Methods

Study population

Data were analyzed from patients enrolled in the Canadian Spine Outcomes and Research Network (CSORN) national registry. CSORN consists of a group of over 50 neurosurgical and orthopedic spine surgeons from 18 tertiary-care academic and nonacademic hospitals across Canada. Each site prospectively collects data on patients undergoing surgical treatment for spinal conditions. For the current project, data were examined from those enrolled in the registry between 2015 and 2018 for which the principal surgical pathology was either LSS, DS, DH or DDD. To standardize pathology classification, the CSORN steering committee (6 spine surgeons) provided an operational definition for surgeons to "Indicate the SINGLE principal pathology for which the surgical technique was chosen." This education and registry form was supplied to all surgeon investigators, who were instructed to choose one principal surgical pathology for each surgical patient.

Ethics approval was obtained from each participating institution and all patient participants consented to be enrolled in the registry. The analyses used data from presurgical questionnaires completed within 6 months of the surgical date and follow-up questionnaires completed at 12 months (+/- 6 weeks) postsurgery were included. Patients who completed questionnnaires outside of these time windows were excluded. Patients were contacted for follow-up based on their preference: phone interview, in-person meeting, regular mail or secure web-based portal. To minimize attrition, patients were informed prior to surgery that follow-up interviews would be conducted at specific time points. They were reminded of this process prior to discharge. Those requesting phone or web surveys were contacted a minimum of 6 times until survey completion was obtained to reduce loss to follow up.

Patient reported outcome measures (PROMs)

1) The *Oswestry Disability Index* (ODI): The modified ODI is a patient-reported measure developed to assess pain-related disability among individuals with low back pain [19,20]. It is among the most frequently used PROMs in surgical spine populations and has demonstrated reliability and validity [21]. The ODI is scored on a 0 to 100 scale, with higher scores indicating greater pain-related disability. Scores are grouped into the following categories: minimal disability (0−20), moderate disability (21−40), severe disability (41−60), crippled (61−80) and bedbound (81−100).
2) *Back* and 3) *Leg Numeric Pain Rating Scales* (Back and Leg NPRS): Patients were asked to indicate on a 0 (No pain) to 10 (Unbearable pain) point scale their overall level of back or leg pain/discomfort. Scores are grouped into the following categories: mild (0−3), moderate (4−6) and severe (7−10). Numeric rating scales for pain are reliable and valid measures of pain intensity in a variety of chronic pain populations [22−25].

The ODI and Back and Leg NPRS were collected in the pre- and the 12-month postsurgical questionnaires. Analyses focused on the absolute change (12 month score − presurgery score) and percentage change ((absolute change/presurgery score)^{x}100) for these three PROMs. Negative absolute or percentage changes represent improvement.

Patient-perceived outcome: In the 12-month postsurgical questionnaire, patients responded to the question '*Compared to before your surgery, how do you feel now?*'

Response options were '*Much better*', '*Better*', '*Same*', '*Worse*', '*Much worse*'. Data from this question were used as the anchor in analyses to determine minimal clinically important difference (MCID) estimates for PROMs.

Additional variables: Patients enrolled in the registry provided data on socio-demographic (sex and age) and health characteristics in the presurgical questionnaire. Body mass index (BMI) was calculated from self-reported height and weight and categorized as: underweight (<18.5 kg/m^2) or normal (18.5-<25 kg/m^2), overweight (25-<30kg/m^2), and obese (30+kg/m^2) [26]. Data on smoking status were grouped as nonsmoker, nonsmoker − uses other nicotine products, former smoker (quit in last 3 months), and current smoker. A count of the number of comorbidities was derived from yes/no responses to a list of 23 conditions (anemia, asthma, cancer, cerebrovascular disease, chronic pulmonary disease, congestive heart failure, connective tissue disease, dementia, depression, diabetes, diabetes with end organ damage, migraine, high blood pressure, high cholesterol, HIV positive, mild liver disease, moderate to severe liver disease, myocardial infarction, nervous system disorders, osteoarthritis, peripheral vascular disease, renal disease, ulcer disease). Spine symptom duration was categorized as ≤6 months, 6 to 12 months, 1 to 2 years and >2 years. On the day of surgery, the surgeon specified whether a spinal fusion was performed as part of the surgical procedure.

Statistical analyses

All analyses were conducted separately for each of the principal pathology groups (LSS, DS, DH, DDD) considered. For all variables, the rate of missingness across the pathology groups ranged from 0% to 1%, except for smoking status where it ranged from 2% to 4%. Patients were included in the MCID calculations irrespective of socio-demographic and health characteristic data availability; all available data were used.

Descriptive: Analyses were conducted to characterize patient socio-demographic and condition characteristics; means and standard deviations were used for continuous variables, and frequencies and percentages for categorical variables. The distribution of responses on the patient-perceived outcome item was determined. To establish that responses to this item were related to both absolute and percentage change in the PROMs of interest, Spearman's correlation coefficients with 95% confidence intervals were estimated. To serve as a suitable anchor in MCID analyses, a moderate correlation of 0.30 to 0.35 or more has been suggested [13,27].

Estimation of MCID thresholds: A receiver operating curve (ROC) approach was used to determine MCID thresholds for each of the three PROMs for both absolute and percentage change. Two sets of MCIDs were estimated: (1) **benefit** for which response was defined as indicating "Much better" or "Better" on the patient-perceived outcome question which served as the analytical anchor, and (2) **substantial benefit** for which response was defined as indicating "Much better". For each level of benefit, the area under the ROC curve (AUC) was estimated in order to examine the level of accuracy in discriminating responders from nonresponders. AUC values of 0.7 to 0.8 are considered acceptable, 0.8 to 0.9 are considered excellent, and more than 0.9 is outstanding [28].

To identify MCID estimates or thresholds, the maximization of Youden's index was selected as the primary criterion. This index maximizes discrimination while equally weighing sensitivity and specificity. For each MCID estimate, the correct classification rate was estimated as the agreement rate for determining responder status based on the MCID estimate with responses on the anchor question. Sensitivity and specificity were also computed. In supplementary analyses, correct classification rates stratified by baseline presurgical status groups were calculated to examine the effect of presurgical status on MCID performance.

Results

There were 3,414 patients with baseline data collected within 6 months of their surgery in the CSORN registry during our study period. Of these, 396 were excluded because they did not complete their follow-up survey within our required time window (1year ±6 weeks). There were no significant differences between those excluded due to the follow-up window and the analytical sample on any baseline characteristics, or for mean baseline ODI or NPRS scores (Supplementary Table 1). Of the 3,018 patients eligible for our study, 866 did not complete their follow-up survey, resulting in an analytical sample of 2,152 (71.3% of eligible patients). Patients who did not complete their follow-up survey were more likely to be undergoing surgery for disc herniation. Given the known clinical profile of this patient population, it was not unexpected to find lower rates of fusion procedures and higher rates of male sex, younger age and current smokers in this group (Supplementary Table 1). Those not completing their follow-up survey also had statistically higher baseline ODI, back and leg NPRS scores, although these differences were relatively small.

Study participants were 856 LSS patients (39.1% female), 591 patients with DS (64.1% female), 520 DH patients (47.5% female) and 185 patients with DDD (43.8% female) (Table 1). The average age of LSS and DS patients was 65.8 years, with DH and DDD patients being younger on average with mean ages of 46.8 years and 50.9 years, respectively. Mean BMI was similar across pathology groups, near the top end of BMIs in the overweight category, ranging from 28 to 29. The proportion of current smokers was highest in DH patients. Most DS (72.8%) and DDD (79.2%) patients had fusion procedures, while fusions were performed in 31.3% of LSS patients and 21.5% of DH patients.

Presurgery scores on the ODI were similar across pathology groups, falling within the lower end of the severe disability category (41−60) [19,21] on average (Table 2).

Table 1
Patient characteristics by pathology

	Mean (SD) or frequency (%)			
	Lumbar spinal stenosis (n = 856)	Degenerative spondylolisthesis (n = 591)	Disc herniation (n = 520)	Degenerative disc disease (n = 185)
Sex				
Female	335 (39.1%)	379 (64.1%)	247 (47.5%)	81 (43.8%)
Male	521 (60.9%)	212 (35.9%)	273 (52.5%)	104 (56.2%)
Age				
Mean	65.8 (11.6)	65.8 (9.4)	46.8 (14.0)	50.9 (12.8)
<65 years	332 (38.8%)	259 (43.8%)	456 (87.7%)	152 (82.2)
65+ years	524 (61.2%)	332 (56.2%)	64 (12.3%)	33 (17.8%)
BMI				
Mean	29.0 (5.5)	28.8 (5.8)	28.0 (5.7)	28.0 (5.5)
Underweight/normal	178 (21.5%)	146 (25.4%)	155 (31.0%)	51 (28.2%)
Overweight	340 (41.1%)	223 (38.7%)	190 (38.0%)	79 (43.7%)
Obese	310 (37.4%)	207 (35.9%)	155 (31.0%)	51 (28.2%)
Smoking status				
Nonsmoker	706 (83.3%)	499 (85.0%)	392 (76.3%)	147 (79.5%)
Nonsmoker, uses other nicotine products	17 (2.0%)	11 (1.9%)	17 (3.3%)	10 (5.4%)
Former smoker (quit in last 3 months)	37 (4.4%)	24 (4.1%)	17 (3.3%)	6 (3.2%)
Current smoker	88 (10.4%)	53 (9.0%)	88 (17.1%)	22 (11.9%)
Comorbidity count	2.3 (1.8)	2.4 (1.8)	1.4 (1.5)	1.6 (1.6)
Condition duration				
≤6 months	76 (8.9%)	23 (3.9%)	116 (22.4%)	11 (6.0%)
>6 months − 12 months	89 (10.5%)	46 (7.8%)	119 (23.0%)	8 (4.4%)
1−2 years	147 (17.3%)	83 (14.1%)	88 (17.0%)	18 (9.8%)
>2 years	538 (63.3%)	435 (74.1%)	194 (37.5%)	147 (79.9%)
Fusion	266 (31.3%)	427 (72.8%)	112 (21.5%)	145 (79.2%)

Mean changes in ODI scores by 12-months postsurgery ranged from a decrease (improvement) of 18.1 points among LSS patients to a decrease of 27.3 among DH patients. Similarly, the percentage change in scores ranged from decreases (improvements) of 38.9% for LSS patients to 55.1% for DH patients.

Back pain ratings presurgery ranged from an average of 6.3 for DH patients to 7.0 for DDD patients (Table 2). The mean leg pain rating was 6.2 for DDD patients, and was 7.4 or 7.5 for the other patient groups. Mean decreases in back pain ratings 12-months postsurgery ranged from 3.2 (DH) to 4.0 (DS) points. Percentage improvements ranged from approximately 46% to 55%. Greater improvements in leg pain ratings than back pain ratings were observed for all groups, except for DDD patients.

A large majority of patients, 84% to 88% across diagnosis groups, met the criterion for achieving benefit, defined as reporting feeling "Much Better" or "Better" 12-months after surgery (Table 3). The percentage achieving substantial benefit (i.e. reporting "Much Better"), was somewhat

Table 2
Oswestry Disability Index (ODI), back pain rating and leg pain rating scores by pathology

	Mean (SD)			
	Lumbar spinal stenosis	Degenerative spondylolisthesis	Disc herniation	Degenerative disc disease
ODI (out of 100; higher=worse)				
Presurgery	45.6 (15.8)	45.3 (15.0)	48.8 (15.8)	47.6 (15.6)
12 months postsurgery	27.4 (19.5)	23.4 (18.2)	21.6 (18.2)	26.3 (20.4)
Change (12 months-presurgery)	−18.1 (18.3)	−22.0 (17.3)	−27.3 (19.9)	−20.9 (17.1)
Percent change	−38.9% (42.8)	−49.3% (38.2)	−55.1% (38.4)	−46.5% (36.3)
Back pain rating (0−10; higher=worse)				
Presurgery	6.8 (2.4)	6.9 (2.4)	6.3 (2.5)	7.0 (2.2)
Twelve months postsurgery	3.4 (2.7)	2.9 (2.5)	3.1 (2.5)	3.5 (2.5)
Change (12 months-presurgery)	−3.3 (3.0)	−4.0 (2.9)	−3.2 (2.9)	−3.5 (2.6)
Percent change	−46.4% (48.7)	−55.2% (42.8)	−46.8% (61.4)	−48.4% (40.8)
Leg pain rating (0−10; higher=worse)				
Presurgery	7.4 (2.1)	7.5 (2.0)	7.4 (1.9)	6.2 (2.8)
12 months postsurgery	3.5 (3.0)	2.9 (3.0)	2.9 (2.7)	3.2 (2.8)
Change (12 months-presurgery)	−3.9 (3.3)	−4.6 (3.2)	−4.6 (3.1)	−3.0 (3.0)
Percent change	−51.1% (50.6)	−59.9% (40.9)	−59.2% (48.0)	−43.2% (62.1)

Table 3
Distribution of surgical outcome anchor for minimal clinically important difference (MCID) determinations and correlations with scales by pathology

	Lumbar spinal stenosis	Degenerative spondylolisthesis	Disc herniation	Degenerative disc disease
Anchor: "Compared to before your surgery, how do you feel now?" N (%)				
Much better	447 (52.5%)	391 (66.8%)	325 (63.0%)	108 (58.7%)
Better	265 (31.1%)	125 (21.4%)	128 (24.8%)	50 (27.2%)
Same	98 (11.5%)	44 (7.5%)	44 (8.5%)	15 (8.2%)
Worse	29 (3.4%)	16 (2.7%)	16 (3.1%)	9 (4.9%)
Much worse	12 (1.4%)	9 (1.5%)	3 (0.6%)	2 (1.1%)
Spearman's correlation (95% CI) between anchor and:				
Change in ODI score	0.59 (0.55, 0.64)	0.49 (0.43, 0.55)	0.60 (0.54, 0.65)	0.59 (0.49, 0.68)
Percent Change in ODI score	0.65 (0.61, 0.69)	0.56 (0.51, 0.62)	0.69 (0.64, 0.73)	0.68 (0.60, 0.75)
Change in back pain rating	0.48 (0.42, 0.53)	0.45 (0.38, 0.51)	0.44 (0.37, 0.51)	0.57 (0.46,0.66)
Percent change in back pain rating	0.55 (0.50, 0.60)	0.55 (0.49, 0.61)	0.59 (0.53, 0.64)	0.68 (0.60, 0.75)
Change in leg pain rating	0.52 (0.47, 0.57)	0.51 (0.45, 0.57)	0.52 (0.46, 0.58)	0.37 (0.23, 0.49)
Percent change in leg pain rating	0.55 (0.51, 0.60)	0.56 (0.51, 0.62)	0.57 (0.50, 0.62)	0.53 (0.41, 0.63)

Table 4
Pathology-specific surgical minimal clinically important difference (MCID) cutoff estimates for "Benefit"*

	Lumbar spinal stenosis	Degenerative spondylolisthesis	Disc herniation	Degenerative disc disease
Change in ODI				
MCID cut-off	−10.5	−12.2	−16.9	−10.0
AUC (95% CI)	0.84 (0.80, 0.87)	0.84 (0.80, 0.89)	0.87 (0.82, 0.91)	0.90 (0.84, 0.95)
Sensitivity / specificity	71.6% / 83.6%	76.1% / 85.5%	77.2% / 85.0%	83.8% / 80.8%
Correct classification rate	73.9%	78.3%	78.3%	83.3%
Responder rate using developed cut-off	63.0%	69.7%	70.0%	74.0%
% Change in ODI				
MCID cut-off	−29.2%	−25.0%	−32.0%	−11.1%
AUC (95% CI)	0.85 (0.82, 0.88)	0.87 (0.83, 0.90)	0.90 (0.87, 0.93)	0.91 (0.87, 0.96)
Sensitivity / specificity	69.4% / 89.6%	82.2% / 81.2%	82.7% / 90.0%	90.9% / 80.8%
Correct classification rate	72.3%	82.1%	83.6%	89.4%
Responder rate using developed cut-off	59.3%	74.5%	73.9%	80.7%
Change in back pain rating				
MCID cut-off	−2.0	−3.0	−2.0	−3.0
AUC (95% CI)	0.80 (0.76, 0.84)	0.82 (0.78, 0.87)	0.81 (0.76, 0.86)	0.88 (0.82, 0.93)
Sensitivity / specificity	77.8% / 69.1%	76.8% / 79.7%	76.4% / 74.2%	73.7% / 92.3%
Correct classification rate	72.1%	78.9%	77.5%	78.0%
Responder rate using developed cut-off	72.2%	71.7%	72.0%	66.7%
% Change in back pain rating				
MCID cut-off	−25.0%	−44.5%	−40.0%	−25.0%
AUC (95% CI)	0.81 (0.77, 0.85)	0.85 (0.80, 0.90)	0.87 (0.83, 0.91)	0.91 (0.87, 0.95)
Sensitivity / specificity	81.9% / 68.4%	76.4% / 91.2%	74.2% / 93.3%	86.5% / 92.0%
Correct classification rate	79.7%	76.2%	76.5%	87.2%
Responder rate using developed cut-off	73.6%	66.0%	65.7%	75.7%
Change in leg pain rating				
MCID cut-off	−2.0	−3.0	−3.0	−2.0
AUC (95% CI)	0.82 (0.78, 0.85)	0.87 (0.83, 0.90)	0.81 (0.76, 0.87)	0.79 (0.70, 0.87)
Sensitivity / specificity	82.4% / 67.4%	78.1% / 81.2%	79.0% / 72.6%	67.7% / 80.8%
Correct classification rate	80.4%	79.3%	79.0%	76.1%
Responder rate using developed cut-off	75.0%	72.0%	73.4%	67.6%
% Change in leg pain rating				
MCID cut-off	−25.0%	−33.4%	−50.0%	−14.3%
AUC (95% CI)	0.82 (0.78, 0.85)	0.89 (0.86, 0.92)	0.84 (0.79, 0.89)	0.84 (0.76, 0.91)
Sensitivity / specificity	81.3% / 71.8%	81.0% / 85.3%	74.5% / 83.9%	83.0% / 73.1%
Correct classification rate	79.8%	80.3%	75.6%	80.3%
Responder rate using developed cut-off	72.9%	71.6%	67.1%	73.6%

* Benefit: ROC analyses use anchor cut points "Much Better, Better" vs. "Same, Worse, Much Worse."

Table 5
Pathology-specific surgical minimal clinically important difference (MCID) cutoff estimates for "Substantial Benefit"*

	Lumbar spinal stenosis	Degenerative spondylolisthesis	Disc herniation	Degenerative disc disease
Change in ODI				
MCID cut-off	−13.8	−22.0	−22.0	−16.0
AUC (95% CI)	0.82 (0.79, 0.85)	0.79 (0.75, 0.83)	0.85 (0.81, 0.88)	0.82 (0.76, 0.88)
Sensitivity / specificity	81.1% / 69.1%	65.2% / 78.7%	80.6% / 71.0%	84.6% / 69.7%
Correct classification rate	76.0%	71.2%	78.1%	78.9%
Responder rate using developed cut-off	57.7%	52.2%	62.9%	61.9%
% Change in ODI				
MCID cut-off	−43.7%	−45.4%	−60.6%	−40.0%
AUC (95% CI)	0.86 (0.84, 0.89)	0.84 (0.80, 0.87)	0.91 (0.88, 0.93)	0.88 (0.83, 0.93)
Sensitivity / specificity	73.3% / 84.1%	75.5% / 80.3%	72.5% / 90.7%	82.7% / 78.9%
Correct classification rate	78.4%	77.1%	79.1%	81.1%
Responder rate using developed cut-off	45.9%	57.1%	49.6%	56.4%
Change in back pain rating				
MCID cut-off	−4.0	−4.0	−4.0	−4.0
AUC (95% CI)	0.75 (0.72, 0.78)	0.76 (0.72, 0.80)	0.75 (0.70, 0.79)	0.81 (0.74, 0.87)
Sensitivity / specificity	64.6% / 73.3%	74.7% / 69.6%	60.7% / 76.6%	73.8% / 78.7%
Correct classification rate	71.1%	75.1%	70.6%	78.6%
Responder rate using developed cut-off	50.4%	62.2%	51.1%	55.7%
% Change in back pain rating				
MCID cut-off	−57.1%	−55.7%	−60.0%	−44.4%
AUC (95% CI)	0.80 (0.77, 0.83)	0.82 (0.79, 0.86)	0.84 (0.80, 0.87)	0.88 (0.83, 0.93)
Sensitivity / specificity	71.5% / 78.0%	78.5% / 75.9%	69.7% / 84.2%	87.7% / 79.7%
Correct classification rate	74.6%	76.8%	75.1%	84.4%
Responder rate using developed cut-off	47.9%	58.6%	49.6%	59.7%
Change in leg pain rating				
MCID cut-off	−5.0	−5.0	−4.0	−3.0
AUC (95% CI)	0.78 (0.75, 0.81)	0.79 (0.76, 0.83)	0.80 (0.76, 0.84)	0.69 (0.61, 0.76)
Sensitivity / specificity	67.9% / 77.0%	71.5% / 76.8%	81.7% / 67.0%	63.9% / 67.1%
Correct classification rate	73.6%	74.8%	77.1%	71.7%
Responder rate using developed cut-off	49.0%	58.1%	64.8%	59.0%
% Change in leg pain rating				
MCID cut-off	−60.0%	−62.5%	−66.6%	−42.9%
AUC (95% CI)	0.80 (0.77, 0.83)	0.82 (0.79, 0.86)	0.83 (0.79, 0.87)	0.78 (0.72, 0.85)
Sensitivity / specificity	74.9% / 75.6%	75.7% / 78.0%	75.2% / 75.9%	77.3% / 68.4%
Correct classification rate	75.2%	76.4%	75.4%	72.3%
Responder rate using developed cut-off	50.9%	57.6%	56.1%	56.3%

* Substantial benefit: ROC analyses use anchor cut points "Much Better" vs. "Better, Same, Worse, Much Worse."

higher for DS (67%) and DH (63%) patients than for DDD (59%) and LSS (53%) patients. Use of responses to this question as the anchor in ROC analyses was supported by moderate to high correlations (range: 0.37-0.69) with each of the three PROMs, for both absolute and percentage change (Table 3).

Table 4 presents the estimates for achieving benefit for each pathology group, for each of the PROMs for absolute and percentage change. MCIDs for ODI absolute change ranged from -10.0 (DDD) to -16.9 (DH). MCIDs for back and leg NPRS absolute change were −2 to −3 for each group. MCIDs for percentage change varied by PROM and pathology group, ranging from −11.1% (ODI for DDD) to −50.0% (leg NPRS for DH). Correct classification rates for all MCID thresholds ranged from 72.1% to 89.4%, with most estimates of sensitivity and specificity falling within this approximate range. AUC values also supported excellent discrimination of responders (patients achieving benefit based on the developed MCID thresholds) and nonresponders. Responder rates also varied by pathology and PROM, ranging from 59.3% (LSS) to 80.7% (DDD) for ODI percentage change. There was less variation in responder rates across groups for the other PROMs, which were mostly on the order of about 70%.

Table 5 contains data analogous to Table 4, but for substantial benefit. These ranged from −13.8 (LSS) to −22.0 (DS, DH) for ODI absolute change and from −40.0% (DDD) to −60.6% (DH) for ODI percentage change. MCID for back NPRS absolute change was −4.0 for all groups and ranged from −44.4% (DDD) to −60.0% (DH) for percentage change. MCID for leg NPRS ranged from −3.0 (DDD) to −5.0 (LSS, DS) and from −42.9% (DDD) to −66.6% (DH) for absolute change and percentage change, respectively. Across pathology groups and PROMs, correct classification rates ranged from 71.2% to 84.4%, with AUC values indicating acceptable to excellent discrimination of responders and nonresponders. Responder rates ranged from 45.9% (LSS) for ODI percentage change to 64.8% (DH) for leg NPRS absolute change.

Relatively few participants fell into the lowest categories of presurgical symptom severity on the ODI, back or leg NPRS across pathology groups (Supplementary Table 1). Correct classification rates for MCIDs according to baseline severity categories were notably lowest for patients in these groups (Supplementary Tables 2 and 3). Improvements in classification accuracy were generally evident for these patients with milder presurgery symptoms when comparing MCIDs for percent change versus absolute change on a given PROM. For example, for DS patients with mild disability on the ODI presurgery, the MCID for benefit based on absolute change had a correct classification rate of 52.9%, but was 76.5% using percent change. There were exceptions to this improvement in accuracy for percent change MCIDs, particularly for the NPRS scales. For those in the highest ODI disability categories (eg, bedbound or crippled) classification accuracy for substantial benefit was also higher for percentage change than absolute change MCIDs.

Discussion

Our study provides minimal clinically important difference (MCID) estimates for the ODI and back and leg NPRS based on four common degenerative lumbar spine surgery cohorts with high classification accuracy. Significant variability in our MCID estimates for these different cohorts was notable across PROMs and evident for both absolute change and percentage change thresholds. These findings suggest that surgical outcome assessment should explicitly consider the lumbar pathology responsible for the procedure in order to improve the accuracy of outcome assessment.

Our MCID thresholds can be used in other patient populations as part of retrospective or prospective research, for example in registry or cohort studies, as well as in clinical trials. Our MCIDs are lumbar pathology and PROM-specific and researchers should apply them in this manner. At an individual patient level, a change, or percentage change, equal to or exceeding the MCID can be used to identify important improvement. MCID thresholds allow for the clinical meaningfulness of changes experienced over time to be assessed [29]. They can be used to facilitate the comparison of outcomes across groups (e.g. treatment effects) via comparison of group-specific success or responder rates [29]. Such rates are calculated as the percentage of patients achieving an improvement greater than or equal to the MCID. Knowledge of responder rates can importantly inform patient expectations and aid patient and provider decision-making, and in addition be used to compare outcomes of new or modified interventions with current interventions. The ability to compare responder rates for "benefit" vs. "surgical benefit" would likely be additionally informative in this regard.

While there are comparable MCID estimates to ours for absolute change in the literature, these were derived in combined lumbar spine pathology populations and their specific applicability to individual pathology groups is not clear. The most commonly used ODI change thresholds have been decreases of 10 [30] or 12.8 [4] points. Our estimates for "benefit" in LSS, DS and DDD were similar to these at 10.5, 12.2 and 10.0 respectively. Our MCID estimate for DH patients was higher than that for the other patient groups, requiring a decrease of 16.9 points on the ODI for 'benefit' and 22.0 for "substantial benefit". Solberg et al. [31] reported a relatively high threshold of 20 points for DH patients in a Norwegian registry study. Our MCID estimates of "benefit" for absolute change in back NPRS were consistent with Austevoll et al. [6] for LSS and DS, requiring decreases of 2 and 3 points, respectively. Our estimates for leg NPRS were analogous to back NPRS for these groups, whereas Austevoll [6] reported a 3-point change for both. Most literature estimates for absolute change on NPRS are on the order of 2 to 3 points [4,7]. The greater similarity of findings for NPRS in the literature may be due in part to their limited range of possible scores; however, we found several rather low estimates of classification accuracy for those with presurgical mild symptoms on NPRS using absolute change scores. This finding suggests that caution is warranted when absolute change score criteria are used to assess outcomes in patients with milder presurgery symptoms.

In order to account for the influence of presurgical symptom severity on surgical outcomes, MCIDs for percent change have been suggested as preferable to those for absolute change, although this has received relatively little attention in the lumbar spine literature. Ostelo et al. [32] concluded that a change of at least 30% from baseline may be clinically meaningful for low back pain patients across a number of PROMs based on literature reviews and expert recommendations. Asher et al. [33] compared the performance of published MCIDs to a 30% threshold on the ODI and NPRS for 12-month surgical outcomes. They concluded that a 30% reduction predicted satisfaction either better than, or as well as, absolute change MCIDs across surgical procedure types for combined lumbar pathologies. Using 30% appeared particularly favorable for those in the lowest and highest ODI categories and the lowest pain categories. We generally found improvements in classification accuracy for these groups as well. To the best of our knowledge, there have only been 2 published studies [6,7] that have statistically derived surgical success criteria for percentage change in PROMs. Glassman et al. [7] calculated percent change criteria for substantial benefit in a group of lumbar spine degenerative surgical pathologies and found changes of 36.8% for ODI, 41.4% for back NPRS and 38.8% for leg NPRS. In 2019, Austevoll et al. [6] used Norwegian registry data and found that thresholds were the same for LSS and DS: 30% for the ODI, 33% for back NPRS and 40% for leg NPRS. In their ROC-based analyses, they used an anchor that defined surgical success as "completely recovered" or "much improved", which seems

the most comparable to our definition for substantial benefit ("much better"). However, our estimates for LSS and DS did vary somewhat, and those for substantial benefit were higher than Austevoll's [6] (LSS, DS: ODI −43.7%, 45.4%; Back NPRS −57.1%, 55.7%; Leg NPRS −60.0%, 62.5%). Estimates for benefit were similar for the ODI, but different for the NPRSs (LSS, DS: ODI −29.2%, 25.0%; Back NPRS −25.0%, 44.5%; Leg NPRS −25.0%, 33.4%). The reasons for these differences may be related to country-specific differences in patient characteristics, expectations, or surgical practices [34,35]. Fusion rates for LSS and DS were 6% and 48% in their study compared to 31% and 73%, respectively, in ours. Notably similar to our findings, they also found percentage change estimates to perform better than absolute change for those with low and high baseline scores. We are not aware of any studies that have evaluated surgical success criteria based on percentage change in PROMs for DH and DDD populations separately.

Our finding of variability in MCIDs across lumbar pathology groups suggests that appropriate outcome assessment requires explicit consideration of the pathology responsible for the surgical procedure. Failure to consider this may result in patient misclassification. For example, using the ODI, we determined a responder rate of 63% for LSS patients achieving benefit based on the LSS-specific MCID. If the DS-specific MCID had been used instead, the LSS responder rate would have been reported as 58%. Conversely, the responder rate for DS patients using the DS-specific MCID was 70% and would have been reported as 73% if the LSS-derived MCID had been used. The degree of misclassification is even greater when considering MCIDs for substantial benefit, due to greater differences in these MCID estimates across pathology groups. For example, using the LSS-specific MCID estimate for substantial benefit yielded a responder rate of 58% in LSS patients, that would be incorrectly reported as 44% if the DS-specific estimate was used.

As an alternative to PROM-specific MCID thresholds, a 30% reduction [32,33] in disability or pain has been suggested as a threshold for clinically relevant improvement after lumbar spine surgery. Several of our estimates were similar to this. For example, our ODI MCIDs for LSS and DH were 29.2% and 32.0%, respectively. However, using 30% across all pathologies and PROMs would have also resulted in some misclassification. For example, we estimated responder rates of 73.6% and 66.0% for LSS and DS patients based on pathology-specific thresholds for percent change in the back NPRS. These responder rates would have been 67.9% and 77.5%, respectively, if a 30% threshold for improvement was used. The difference in responder rates is also quite notable for DDD patients based on the ODI percentage change threshold. The responder rate would have been 65.2% if a 30% threshold was used, while we determined a responder rate of 80.7%.

The identified variability in MCIDs across pathology groups likely stems from multiple factors. Numerically, the estimates make intuitive sense given the relative distributions of pre- and postsurgical PROM scores across groups. For example, given the larger absolute and percentage change in ODI scores, on average, in DH patients compared to the other patient groups, it is not surprising that the DH-derived MCIDs are also greater. Also, differences in patient age, associated comorbidities and perceptions of surgical impact, and their influence on PROMs and anchor question responses, likely reflect the variability in the dominant symptoms across pathology groups and their respective changes due to surgery (eg, intermittent neurogenic claudication in DS/LSS, constant radiculopathy in DH and mechanical back pain in those with DDD). Further, patients' expectations of outcome, which are influenced by perioperative education on likely prognosis conveyed by the surgeon, also may influence PROMs. Perioperative education will accordingly vary based on known differences in surgical response across pathologies, with more favorable surgical response often conveyed to patients undergoing surgery for DH and DS vs. LSS and DDD. Consequently, this may contribute to the seemingly paradoxical finding that DH and DS patients require a larger change to perceive surgical benefit. It should be noted that expectations in the lumbar spine population are highly variable amongst patients and between patients and surgeons. [2,36]

Strengths and limitations

Strengths of our study include the use of a comprehensive registry with broad Canadian coverage. Further, we utilized a narrow time window of 12 months +/− 6 weeks postsurgery to ensure we were capturing a consistent recovery time point. Future work may wish to consider an additional recovery window of 2 years postsurgery. However, Alying et al [37] demonstrated using CSORN registry data that there was no significant difference in PROMs at 1- vs. 2 -years postsurgery for degenerative lumbar pathologies. The importance of our separate consideration of homogenous patient pathology groups is clearly evident in the variability in our findings for these groups. Our inclusion of DH and DDD patients in our study is also an important addition to the literature, as pathology-specific estimates for these groups are particularly limited. Variability in our findings with those reported for other jurisdictions may reflect differences in patient or health system characteristics or surgical practices. Consideration of the characteristics of our sample in terms of available patient and surgical factors may aid in assessing the relative generalizability of our findings to other specific populations. While the generalizability of our derived MCID estimates to other jurisdictions warrants consideration, our findings nonetheless strongly support the need to consider principal surgical pathology explicitly in assessing outcomes in lumbar surgery. To fully assess whether a patient's surgery is a "success" may require considering

achievement of benefit or substantial benefit on more than one PROM, including ones we did not consider, and potentially those specific to a given patient's surgical outcome goals. This may also necessitate considering objective functional measures. Nevertheless, estimates of the performance of our MCID estimates were highly supportive overall. Our estimates were dependent on the anchor used in ROC analyses - a single-item of global patient-perceived outcome. Such an item requires a retrospective assessment of presurgical status by the patient and may be influenced by current status [38]. While the validity of such an assessment for the specific study populations has not been determined, use of a patient-perceived outcome is recommended [16] and an item similar to that used in our work has been used in several other patient populations [29,39–41]. Further, the FDA considers such an assessment a suitable anchor for defining PROM thresholds [42] and the IMMPACT recommendations for chronic pain clinical trials support the use of a global patient assessment of outcome [29]. Use of our selected anchor was supported by a high correlation with the PROMs studied [13]. Our global measure of patient-perceived outcome had 5 possible response options, while a 7-level response scale has also been used in the literature [4,5,8,9,31]. In using a 5-level anchor, we are essentially missing categories of minimal improvement or worsening such as "somewhat" (or "minimally" or "slightly") better or worse. However, Dworkin et al. [29] recommended not combining these "minimal" response options with the other categories of global improvement as it is unclear how important these minimal changes are to patients. Future work should consider estimating analogous MCID thresholds in different study samples, as well as the use of different anchors, to compare the stability and reproducibility of our findings. Our thresholds for DDD were notably smaller than for the other groups, particularly for ODI percentage change (−11.1%), suggesting these patients require a smaller relative change to feel "better". Given the smaller sample size of our DDD patient group, and that these patients are rarely considered for surgery in the Canadian healthcare system, further validation work in this population is warranted. Further, additional work should consider how patient characteristics such as sex influence MCID estimates.

Our study establishes a comprehensive set of surgical responder criteria for the most common pathologies in surgical lumbar spine patients. Our findings suggest that the use of generic MCID thresholds across lumbar spine pathologies is not recommended; furthermore, it is important to consider presurgical symptom status in assessing postsurgical outcomes. In particular, for patients with relatively low or high presurgery PROM scores, MCIDs based on percentage change, rather than absolute change, appear generally preferable. Our findings may be applicable in clinical and research settings, and establish an important benchmark for utilization of principal pathology specific MCIDs in future surgical prognostic work.

Declarations of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Perruccio is supported by an award from Arthritis Society Canada (STAR-20-0000000012). The authors have no financial or personal relationships that could inappropriately influence this work.

Supplementary materials

Supplementary material associated with this article can be found in the online version at https://doi.org/10.1016/j.spinee.2023.05.001.

References

[1] 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):541–6.

[2] Rampersaud YR, Canizares M, Perruccio AV, Abraham E, Bailey CS, Christie SD, et al. Fulfillment of patient expectations after spine surgery is critical to patient satisfaction: A cohort study of spine surgery patients. Neurosurgery 2022;91(1):173–81.

[3] Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials 1989;10(4):407–15.

[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 2008;8(6):968–74.

[5] Parker SL, Mendenhall SK, Shau D, Adogwa O, Cheng JS, Anderson WN, et al. Determination of minimum clinically important difference in pain, disability, and quality of life after extension of fusion for adjacent-segment disease. J Neurosurg Spine 2012;16(1):61–7.

[6] Austevoll IM, Gjestad R, Grotle M, Solberg T, Brox JI, Hermansen E, et al. Follow-up score, change score or percentage change score for determining clinical important outcome following surgery? An observational study from the Norwegian registry for Spine surgery evaluating patient reported outcome measures in lumbar spinal stenosis and lumbar degenerative spondylolisthesis. BMC Musculoskelet Disord 2019;20(1):31.

[7] Glassman SD, Copay AG, Berven SH, Polly DW, Subach BR, Carreon LY. Defining substantial clinical benefit following lumbar spine arthrodesis. J Bone Joint Surg Am 2008;90(9):1839–47.

[8] 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 spondylolisthesis. J Neurosurg Spine 2011;14(5):598–604.

[9] 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):471–8.

[10] Fekete TF, Haschtmann D, Kleinstuck FS, Porchet F, Jeszenszky D, Mannion AF. What level of pain are patients happy to live with after surgery for lumbar degenerative disorders? Spine J 2016;16(Suppl. 4):S12–8.
[11] Carreon LY, Bratcher KR, Canan CE, Burke LO, Djurasovic M, Glassman SD. Differentiating minimum clinically important difference for primary and revision lumbar fusion surgeries. J Neurosurg Spine 2013;18(1):102–6.
[12] Wright A, Hannon J, Hegedus EJ, Kavchak AE. Clinimetrics corner: A closer look at the minimal clinically important difference (MCID). J Man Manip Ther 2012;20(3):160–6.
[13] Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61(2):102–9.
[14] McGlothlin AE, Lewis RJ. Minimal clinically important difference: Defining what really matters to patients. JAMA 2014;312 (13):1342–3.
[15] Turner D, Schunemann HJ, Griffith LE, Beaton DE, Griffiths AM, Critch JN, et al. The minimal detectable change cannot reliably replace the minimal important difference. J Clin Epidemiol 2010;63 (1):28–36.
[16] Devji T, Carrasco-Labra A, Qasim A, Phillips M, Johnston BC, Devasenapathy N, et al. Evaluating the credibility of anchor based estimates of minimal important differences for patient reported outcomes: Instrument development and reliability study. BMJ 2020;369: m1714.
[17] Little DG, MacDonald D. The use of the percentage change in Oswestry Disability Index score as an outcome measure in lumbar spinal surgery. Spine (Phila Pa 1976) 1994;19(19):2139–43.
[18] Lauridsen HH, Hartvigsen J, Manniche C, Korsholm L, Grunnet-Nilsson N. Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain patients. BMC Musculoskelet Disord 2006;7:82.
[19] Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiotherapy 1980;66(8):271–3.
[20] Fritz JM, Irrgang JJ. A comparison of a modified Oswestry low back pain disability questionnaire and the Quebec back pain disability scale. Phys Ther 2001;81(2):776–88.
[21] Fairbank JC, Pynsent PB. The Oswestry Disability Index. Spine (Phila Pa 1976) 2000;25(22):2940–52. discussion 2952.
[22] Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis Care Res (Hoboken) 2011;63(Suppl. 11):S240–52.
[23] Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. Spine (Phila Pa 1976) 2005;30(11):1331–4.
[24] Hjermstad MJ, Fayers PM, Haugen DF, Caraceni A, Hanks GW, Loge JH, et al. European palliative care research C: Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults—a systematic literature review. J Pain Symptom Manage 2011;41(6):1073–93.
[25] Karcioglu O, Topacoglu H, Dikme O, Dikme O. A systematic review of the pain scales in adults: Which to use? Am J Emerg Med 2018;36 (4):707–14.
[26] Health Canada. Canadian Guidelines for Body Weight Classification in Adults. Ottawa: Minister of Public Works and Government Services Canada; 2003.
[27] Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Earlbaum Associates; 1988.
[28] Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol 2010;5(9):1315–6.
[29] Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain 2008;9(2):105–21.
[30] 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.
[31] Solberg T, Johnsen LG, Nygaard OP, Grotle M. Can we define success criteria for lumbar disc surgery?: Estimates for a substantial amount of improvement in core outcome measures. Acta Orthop 2013;84(2):196–201.
[32] 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):90–4.
[33] Asher AM, Oleisky ER, Pennings JS, Khan I, Sivaganesan A, Devin CJ, et al. Archer KR: Measuring clinically relevant improvement after lumbar spine surgery—is it time for something new? Spine J 2020;20(6):847–56.
[34] Lagerback T, Fritzell P, Hagg O, Nordvall D, Lonne G, Solberg TK, et al. Effectiveness of surgery for sciatica with disc herniation is not substantially affected by differences in surgical incidences among three countries: Results from the Danish, Swedish and Norwegian spine registries. Eur. Spine J. 2019;28(11):2562–71.
[35] Lonne G, Fritzell P, Hagg O, Nordvall D, Gerdhem P, Lagerback T, et al. Lumbar spinal stenosis: Comparison of surgical practice variation and clinical outcome in three national spine registries. Spine J 2019;19(1):41–9.
[36] Aoude A, Litowski M, Aldebeyan S, Fisher C, Hall H, Manson N, et al. A comparison of patient and surgeon expectations of spine surgical outcomes. Global Spine J 2021;11(3):331–7.
[37] Ayling OGS, Ailon T, McIntosh G, Soroceanu A, Hall H, Nataraj A, et al. Clinical outcomes research in spine surgery: What are appropriate follow-up times? J Neurosurg Spine 2018;30(3):397–404.
[38] Grovle L, Haugen AJ, Hasvik E, Natvig B, Brox JI, Grotle M. Patients' ratings of global perceived change during 2 years were strongly influenced by the current health status. J Clin Epidemiol 2014;67(5):508–15.
[39] Farrar JT, Young Jr. JP, LaMoreaux L, Werth JL, Poole MR. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 2001;94(2):149–58.
[40] Escobar A, Garcia Perez L, Herrera-Espineira C, Aizpuru F, Sarasqueta C, Gonzalez Saenz de Tejada M, et al. Total knee replacement; minimal clinically important differences and responders. Osteoarthritis Cartilage 2013;21(12):2006–12.
[41] Cella D, Hahn EA, Dineen K. Meaningful change in cancer-specific quality of life scores: Differences between improvement and worsening. Qual. Life Res. 2002;11(3):207–21.
[42] McLeod LD, Coon CD, Martin SA, Fehnel SE, Hays RD. Interpreting patient-reported outcome results: US FDA guidance and emerging methods. Expert Rev Pharmacoecon Outcomes Res 2011;11(2):163–9.