The Spine Journal 19 (2019) 1319−1323

THE SPINE JOURNAL

Clinical Study

The minimum detectable measurement difference for the Scoliosis Research Society-22r in adolescent idiopathic scoliosis: a comparison with the minimum clinically important difference

Michael P. Kelly, MD, MSc[a,*], Lawrence G. Lenke, MD[b], Paul D. Sponseller, MD[c], Joshua M. Pahys, MD[d], Tracey P. Bastrom, MA[e], Baron S. Lonner, MD[f], Mark F. Abel, MD[g]

[a] *Washington University School of Medicine, Saint Louis, MO, USA*
[b] *The Spine Hospital, Columbia College of Physicians and Surgeons, 5141 Broadway, New York, NY 10034, USA*
[c] *Division of Pediatric Orthopaedics, All Children's Hospital at Johns Hopkins, 1800 Orleans St., Baltimore, MD 21287, USA*
[d] *Shriners' Hospital for Children, 3551 N Broad St, Philadelphia, PA 19140, USA*
[e] *Department of Orthopedic Surgery, Rady Children's Hospital, 3020 Children's Way, San Diego, CA 92123, USA*
[f] *Department of Orthopaedic Surgery, Mount Sinai Hospital, 1 Gustave L. Levy Pl, New York, NY 10029, USA*
[g] *University of Virginia School of Medicine, 1215 Lee St, Charlottesville, VA 22908, USA*
Received 23 October 2018; revised 3 April 2019; accepted 9 April 2019

Abstract

BACKGROUND CONTEXT: The minimal clinically important difference (MCID) is the smallest change in an outcomes instrument deemed relevant to a patient. MCID values proposed in spine research are limited by poor discriminative abilities to accurately classify patients as "improved" or "not improved." Furthermore, the MCID should not compare relative effectiveness between two groups of patients, though it is frequently used for this. The minimum detectable measurement difference (MDMD) is an alternative to the MCID in outcomes research. The MDMD must be greater than the MCID for the latter to be of value and the MDMD can compare change between groups.
PURPOSE: The purpose of this study was to determine the MDMD for the Scoliosis Research Society-22r (SRS-22r) in adolescent idiopathic scoliosis (AIS) patients treated with surgery.
STUDY DESIGN: Retrospective cohort study from multi-center registry.
PATIENT SAMPLE: Patients treated surgically for AIS.
OUTCOME MEASURES: Self-reported SRS-22r.
METHODS: An observational cohort of surgically treated AIS patients was queried for patients with complete baseline, 1-year, and 2-year SRS-22r data. The MDMD was calculated for SRS-22r domain and subscores. Effect size (ES) and standardized response mean were calculated to measure responsiveness of the SRS-22r to change. MDMD values were compared with MCID values. Research grants were received from DePuy Synthes Spine, EOS imaging, K2M, Medtronic, NuVasive, and Zimmer Biomet to Setting Scoliosis Straight Foundation.
RESULTS: One thousand two hundred and eighty-one AIS patients (1,034 female, 247 male, mean age 14.6 years) were analyzed. MDMD values were between 0.23 and 0.31. SRS-Pain MDMD was 0.3, greater than the MCID of 0.2. SRS-Activity MDMD was 0.24, greater than the

FDA device/drug status: Not applicable.

Author disclosures: ***MPK:*** Grant: Depuy Synthes Spine to SSSF (B). ***LGL:*** Nothing to disclose. ***PDS:*** Grant: Depuy Synthes Spine to SSSF; Royalties: Depuy Synthes spine, Globus (F). ***JMP:*** Grant: Depuy Synthes Spine to SSSF; Consulting: DePuy Synthes (B), NuVasive (B), Zimmer Biomet (B). ***TPB:*** Grant: Depuy Synthes Spine to SSSF. ***BSL:*** Grant: Depuy Synthes Spine to SSSF (D), John and Marcella Fox Fund (B), OREF (C); Royalties: Depuy Synthes (G); Stock Ownership: Paradign Spine (D), Spine Search (E); Private Investments: Paradigm Spine (E); Consulting Depuy Synthes (D), Zimmer Biomet (B), Apifix (B), Unyg Align (B); Speaking and/or Teaching Arrangements: Depuy Synthes (C), K2M (C); Board of Directors: Spine Search; Scientific Advisory Board/Other Office: Depuy Synthes. ***MFA:*** Grant: Depuy Synthes Spine to SSSF; Research Support (Investigator Salary, Staff): Harms Study Group & SSS LLC.

* Corresponding author. Department of Orthopedic Surgery, Washington University School of Medicine, 660 South Euclid Avenue, Box 8233, Saint Louis, MO, USA. Tel: 314-747-2511; fax: 314-747-2500.
E-mail address: kellymi@wudosis.wustl.edu (M.P. Kelly).

https://doi.org/10.1016/j.spinee.2019.04.008
1529-9430/

MCID of 0.08. SRS-self-image MDMD was 0.3, less than the MCID of 0.98. Sixty-four percent of those with baseline SRS-self-image>4.0 improved MDMD or more, whereas only 14% improved beyond the MCID. ES and standardized response mean were highest for subscore and self-image.
CONCLUSIONS: The MDMD can compare the relevance of change in SRS-22r scores between groups of AIS patients. SRS-pain and SRS-activity MDMD values are greater than the MCID and should serve as the threshold for clinically relevant improvement. MDMD may help evaluate change in patients with baseline self-image>4.0.

Keywords: Adolescent idiopathic scoliosis; Minimum clinically important difference; Patient reported outcomes; SRS-22r; Comparative effectiveness; Minimum detectable change

Introduction

The minimum clinically important difference (MCID) is the smallest change of clinical relevance in a health-related quality of life (HRQOL) questionnaire and may be an essential component of patient-reported outcomes research. It is a measure of whether a patient feels improved, relative to the change in outcomes scores. The MCID is often calculated by distribution-based or anchor-based methods [1]. With distribution-based techniques, the MCID is a multiple of the standard error of measurement (SEM) or some other fixed value related to the dataset (eg, standard deviation). Anchor-based methods rely on an external question (the anchor) to assess the clinical importance of change, as assessed by the patient. For example, the patient is asked “Are you improved? Yes or no?” The response is used to select a value that maximizes the sensitivity and specificity of the MCID, thus offering a more accurate estimate of MCID. Neither method is perfect and each has unique limitations. A common misconception regarding MCID is that it can be used to evaluate the clinical relevance of differences between treatment groups [2].

The minimum detectable measurement difference (MDMD) is a complement to the MCID [3,4]. The MDMD is the smallest change detectable outside the measurement error of the outcomes instrument. Therefore, the MCID must be larger than the MDMD. In some situations, the MCID may be so large that certain groups cannot improve to achieve the MCID. For example, the MCID proposed for the SRS-22 self-image domain is 0.98 [5]. As a result, analyses often exclude patients with baseline self-image scores equal to or greater than 4.0, making the assumption that these patients cannot have relevant improvements [6]. Finally, the MDMD offers a threshold for significant differences between groups, allowing for comparisons of change between treatment groups (eg, if the difference between groups is less than the MDMD, then no difference exists). Comparative effectiveness studies, such as tethering vs fusion, will benefit from the use of this MDMD concept.

The MDMD for the SRS-22r in AIS is not known. The SRS-22 is a valid and reliable outcomes instrument in adolescent idiopathic scoliosis (AIS) [7−9]. Carreon et al. used anchor-based calculations to propose an MCID of 0.2 for Pain, 0.08 for activity, and 0.98 for self-image in patients with AIS using the SRS-22 instrument [5]. The purpose of this study was to calculate the MDMD for the SRS-22r and to investigate the effect of gender on the MDMD in AIS. A secondary goal of the study was to assess the responsiveness of the SRS-22r to change by calculating the effect size (ES) and standardized response mean (SRM). These are two measures of the ability of the SRS-22r to detect change in response to treatment, with larger values implying a greater sensitivity. Knowledge of the MDMD, ES, and SRM will provide for a better understanding of reliability of SRS-22r scores and allow for evaluation of between group differences in comparative effectiveness research.

Materials and methods

A multicenter prospective, observational cohort of patients with surgical AIS patients was queried from years 2002 to 2013. Institutional Review Board approval was obtained and guardian/parent consent obtained for enrollment. Patients were excluded from this analysis if any SRS-22r data were missing at baseline, 1 year, and 2 years after surgery. Only patients with complete outcomes data at a minimum of 2 years follow-up were included. Demographic data collected included age at enrollment, gender, standard radiographic measurements, and baseline, 1-year, and 2-year SRS-22r scores. The MDMD at the group level was calculated according to the method described by Spratt [4]. This calculation is a measure of standard errors of measurement of change over 1 and 2 years after surgery. The SEM at the group-level was calculated for the baseline, 1-year, and 2-year values. The MDMD was calculated for the whole cohort and according to gender.

The ES is a measure of responsiveness to change and is equal to mean change in scores divided by the standard deviation of the baseline scores. The SRM is another measure of responsiveness to change and is equal to the mean change in scores divided by the standard deviation of the change in scores. Larger ES and SRM indicate a greater responsiveness to change. The values are interpreted as proposed by Cohen, where a large ES is 0.8 or greater, moderate is 0.5, and small is less than 0.2 [10].

To investigate the effect of using MCID and MDMD as thresholds for relevant change, the proportion of patients meeting MCID and MDMD for SRS-22r pain, self-image, and activity domains were calculated. For self-image, a subset of patients with baseline SRS-self-image scores of 4.0 or greater was selected to examine the effect of MCID (0.98) vs MDMD as a threshold of change.

Continuous data and categorical variables were compared between groups with un-paired *t* tests and Chi-squared tests, respectively. All statistical calculations were performed using SPSS v23.0 (IBM, Chicago, IL, USA) and Microsoft Excel 2016 (Microsoft, Seattle, WA, USA). Statistical significance was defined as p<.05.

Sources of support

This study was supported in part by grants to the Setting Scoliosis Straight Foundation in support of Harms Study Group research from DePuy Synthes Spine, EOS imaging, K2M, Medtronic, NuVasive, and Zimmer Biomet.

Results

Two thousand eight hundred and ninety patients were enrolled and 1,281 (44%) had complete data available for inclusion. There were 1,034 females and 247 males. The average age at enrollment was 14.6 and females were slightly younger (range 10−22, female 14.3±2.1, male 15.9±1.9, p<.0001). There were significantly more Lenke Type 1 curves in the female group, although Lenke 2 curves were more common in the male group ($\chi 2$ (5, N=1,281)=30.56, p<.0001, Table 1).

The group-level MDMD for each SRS-22r domain and subscore lay between 0.23 and 0.31 (Table 2). The MDMD was consistent from at 1 and 2 years postoperatively for all scores. The MDMD was relatively consistent for each gender, both between genders and at 1 and 2-year follow-up. The MDMD for SRS-pain was 0.3, greater than the proposed MCID value of 0.2. The MDMD for SRS-activity was 0.24, greater than the proposed MCID value of 0.08. The MDMD for SRS-self-image was 0.3, less than the proposed MCID value of 0.98. Nearly one-quarter of the cohort (302/1,281, 23.6%) had a baseline SRS-self-image score greater than or equal to 4.0. Only 14% of these patients improved to 5.0 (to meet MCID), whereas 64% improved beyond the MDMD (Table 3).

Table 1
Distribution of lenke curve types by gender

Lenke curve type	Total (N=1,281)	Female (N=1,034)	Male (N=247)
1	**552 (43.1%)**	**462 (44.7%)**	**90 (36.4%)**
2	**272 (21.2%)**	**188 (18.2%)**	**84 (34.0%)**
3	**93 (7.3%)**	**76 (7.4%)**	**17 (6.9%)**
4	**46 (3.6%)**	**38 (3.7%)**	**8 (3.2%)**
5	**196 (15.3%)**	**166 (16.1%)**	**30 (12.1%)**
6	**120 (9.4%)**	**102 (9.9%)**	**18 (7.3%)**

χ^2 (5, N=1,281)=30.56, p<.0001.

Table 2
Minimum detectable measurement difference (MDMD), effect size, and standardized response means at 1-year and 2-years postoperative

SRS-22r domain	MDMD	Effect size	Standardized response mean	Minimum clinically important difference*
Pain				0.2
1 Year	0.30	0.54	0.53	
Female	0.25	0.53	0.52	
Male	0.22	0.63	0.65	
2 Year	0.30	0.51	0.49	
Female	0.22	0.61	0.46	
Male	0.19	0.91	0.64	
Self-image				0.98
1 Year	0.30	1.62	1.51	
Female	0.25	1.59	1.48	
Male	0.22	1.59	1.48	
2 Year	0.31	1.60	1.43	
Female	0.22	2.19	1.43	
Male	0.21	2.04	1.49	
Activity				0.08
1 Year	0.24	0.13	0.11	
Female	0.21	0.10	0.09	
Male	0.19	0.25	0.24	
2 Year	0.24	0.38	0.39	
Female	0.19	0.42	0.35	
Male	0.15	0.54	0.45	
Mental health				N/A
1 Year	0.27	0.43	0.45	
Female	0.23	0.46	0.49	
Male	0.22	0.38	0.38	
2 Year	0.30	0.42	0.39	
Female	0.24	0.24	0.40	
Male	0.22	0.22	0.38	
Subscore				N/A
1 Year	0.23	1.02	1.09	
Female	0.16	1.10	1.16	
Male	0.16	1.09	1.12	
2 Year	0.24	1.13	1.10	
Female	0.16	1.41	1.08	
Male	0.13	1.49	1.21	

MCID, minimal clinically important difference.
* MCID from Carreon et al. [5].

Table 3
Comparison of patients meeting MCID and MDMD thresholds for pain, self-image, and activity

Domain	Percentage of patients meeting threshold value (number of patients)
Pain	
MCID (0.2)	59.8% (766)
MDMD (0.3)	50.9% (652)
Self-image	
MCID (0.98)	57.8% (741)
MDMD (0.3)	85.5% (1095)
Subset of baseline self-image≥4.0	
MCID (0.98)	14.2% (43)
MDMD (0.3)	63.6% (192)
Activity	
MCID (0.08)	51.4% (658)
MDMD (0.24)	39% (499)

Effect size and SRM were largest for self-image at 1-year and 2-year follow-up (ES 1.62/1.60, SRM 1.51/1.43) and subscore (ES 1.02/1.13, SRM 1.09/1.10) indicating that the SRS-22r is most sensitive to change in self-image and subscore. Effect size and SRM were smallest, indicating poor sensitivity to change, for activity and only slightly higher for pain and mental health (Table 2). Effect size and SRM were relatively stable over time and across genders.

Discussion

The MCID is a threshold value for clinical benefit when assessing change in HRQOL scores. In the case of AIS, quantifying "benefit" may be difficult as surgery is performed to prevent deterioration in HRQOL over the long-term. Despite this, investigations into the MCID for AIS using the SRS-22r instrument exist [5]. Subsequent studies have reported the results of AIS surgery, using MCID as a measure of success [6]. This threshold for clinical benefit may be excessive, given the ceiling effects that exist with the instrument and a poor ability to discriminate change for these generally healthy patients. In some cases, the MCID may fall within the error of measurement, which is the MDMD. For a proposed MCID to be useful, it must be greater than the MDMD (the error of measurement of change) [4]. The purpose of this study was to define the MDMD for the SRS-22r instrument domains in a cohort of patients treated for AIS and to compare it to reported MCID values.

We found consistent MDMD values across SRS-22r domains for AIS patients, with values ranging from 0.24 to 0.31. There was essentially no effect of gender on the calculated MDMD and use of general MDMD values for SRS-22r domains is appropriate. The SRS-22r was most sensitive to change in the self-image domain and least sensitive to activity as seen by the large and small ES/SRM, respectively. These results may not be surprising, given the effect of surgery on cosmesis. Conversely, many AIS patients are not disabled before or after surgery. This results in a small change in the activity domain after surgery. That the responsiveness to change for activity is poor has implications for comparisons of fusion and motion-sparing techniques. If the instrument is not able to detect change before and after fusion, then it is unlikely that it will show any difference between the two types of surgery. A more comprehensive assessment of function and motion after surgery is needed to detect differences and significant effects of motion-sparing techniques.

The MDMD for the pain domain was 0.3, which is greater than the MCID value proposed by Carreon et al. using anchor-based estimations, 0.2 [5]. Thus, we propose that the MCID for pain be set at 0.3. With this threshold, there is confidence that a clinical benefit has occurred and the change is outside of the error of the instrument. Similarly, the MDMD for the activity domain was approximately 0.3, whereas the anchor-based value calculated is 0.08 which was less than the minimum detectable change calculated by the authors. The MCID for the activity domain, therefore, should be a minimum of 0.3. The self-image MDMD was also 0.3, which is less than the anchor-based estimation and the MCID should remain set at 0.98, as previously proposed. In this cohort, nearly 25% of the patients had baseline SRS-self-image scores of 4.0 or greater. Evaluation of success in this domain is difficult in this subset of patients, as improvement to a perfect score of five may not be a reasonable expectation [6]. We found that only 14.2% of patients improved to 5.0, although 64% improved some amount greater than the MDMD. This is consistent with the “30% Change Rule,” proposed as an alternative to anchor-based and distribution-based MCID calculations [4]. Spratt has shown that this concept is useful when baseline HRQOL values are high and improvement to MCID or greater is unlikely or impossible. In the case of SRS-self-image scores equal to 4.0 the “30% Change Rule” would define MCID as 0.3 ($5-4=1\times0.3$). When evaluating change in these patients, use of the MDMD value of 0.3 may offer some insight into improvement and the lack thereof. Ceiling effects exist for both SRS-pain and SRS-activity, as AIS is not often debilitating, and smaller ESs and SRMs were calculated. Changes in appearance may be the most noticeable by the patient after AIS surgery, which is seen in the larger calculated ES and SRM.

Carreon et al. analyzed 887 patients treated with surgery for AIS [5]. The MCID was calculated for the SRS-22 pain, activity, and self-image domains using the anchor-based method with receiver operating characteristic curves. The authors present a careful and accurate interpretation of their results. The area under the curve calculated for each of the three domains is less than 0.7. This indicates fair to good accuracy of the proposed MCID when determining a clinically significant change. The importance of improved cosmesis is reflected in the calculated MCID for self-image, 0.98, and the smaller SEM (0.21) and minimum detectable change (0.47) [11]. This is reflected by an MCID that is less than the minimum detectable change and the ceiling effect present within the instrument [12].

The MDMD provides an alternative to the MCID, which may be useful in several situations. First, when the estimated MCID is less than calculated minimum detectable change by the instruments as is the case with the SRS-22r pain and activity domains. In such a case, the MDMD can define “real” change, knowing that this change is also clinically relevant. A second situation where the MDMD may be a useful adjunct to the MCID is when baseline measurements make improvement beyond the MCID impossible, for example an SRS-22r self-image domain score of 4.2. This patient cannot improve beyond 5.0, thus the MCID is an unobtainable goal. The MDMD provides an alternative measure of success in these patients. Finally, the MDMD allows for the comparison of change between two groups, which the MCID does not because the MCID is a “within individual" measure and not a measure between groups [2]. A difference of 0.2 between self-image scores in two comparison groups lies within the MDMD and no difference may exist, despite statistical significance. Finally, the MDMD can serve as a guide for future sample size estimations in comparative effectiveness research.

These data and conclusions have some limitations. We have investigated AIS only and these conclusions may not translate to other pediatric spinal diagnoses, such as syndromic scoliosis or neuromuscular scoliosis. There were a substantial number of patients excluded, caused by missing data. In a study examining effectiveness of a procedure or intervention, this may be an issue. In this case, however, we are examining properties of the instrument and the bias associated with loss is unlikely to affect the conclusions. Furthermore, the SRS instrument scale is small, up to five points, and mean values and standard deviations would be unlikely to change substantially with the addition of lost patients. We do not have a balanced proportion of Lenke curve types, though the distribution of curve types is similar to prior work and it likely reflects the distribution in clinical practice [13]. Perhaps the most important limitation is that MDMD does not confer any information regarding the perception of change by the patient. It is an improvement on MCID in statistical terms only; in other words, we can state whether the change is outside the error of measurement and may be of clinical relevance. Future work investigating the MDMD with an anchor question regarding improvement will clarify this limitation.

Conclusions

MCID is often used, incorrectly, to compare the relevance of between group changes leading to erroneous conclusions of "no difference." The MDMD allows for comparisons of differences between groups and detection of real differences between groups. Thus the MDMD offers a threshold to evaluate differences in future comparative effectiveness studies examining treatment options for AIS, such as fusion and motion-sparing techniques. Finally, in any case where the MCID is used as an outcome measure, it should be larger than the MDMD. If it is not, the MDMD should serve as the threshold for clinically relevant change.

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. The spine journal: official journal of the North American Spine Society 2007;7:541–6.

[2] Chung AS, Copay AG, Olmscheid N, Campbell D, Walker JB, Chutkan N. Minimum Clinically Important Difference: Current Trends in the Spine Literature. Spine 2017;42:1096–105.

[3] Dhawan V, Spratt KF, Pinzur MS, Baumhauer J, Rudicel S, Saltzman CL. Reliability of AOFAS diabetic foot questionnaire in Charcot arthropathy: stability, internal consistency, and measurable difference. Foot & ankle international 2005;26:717–31.

[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 2009;34:1722–31.

[5] Carreon LY, Sanders JO, Diab M, Sucato DJ, Sturm PF, Glassman SD. The minimum clinically important difference in Scoliosis Research Society-22 Appearance, Activity, And Pain domains after surgical correction of adolescent idiopathic scoliosis. Spine 2010;35: 2079–83.

[6] Bennett JT, Samdani AF, Bastrom TP, et al. Factors affecting the outcome in appearance of AIS surgery in terms of the minimal clinically important difference. Eur Spine J 2016.

[7] Parent EC, Hill D, Mahood J, Moreau M, Raso J, Lou E. Discriminative and predictive validity of the scoliosis research society-22 questionnaire in management and curve-severity subgroups of adolescents with idiopathic scoliosis. Spine 2009;34:2450–7.

[8] Asher M, Min Lai S, Burton D, Manna B. The reliability and concurrent validity of the scoliosis research society-22 patient questionnaire for idiopathic scoliosis. Spine 2003;28:63–9.

[9] Asher M, Min Lai S, Burton D, Manna B. Discrimination validity of the scoliosis research society-22 patient questionnaire: relationship to idiopathic scoliosis curve pattern and curve size. Spine 2003;28:74–8.

[10] Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed.. Hillsdale, N.J.: L. Erlbaum Associates; 1988.

[11] Bridwell KH, Shufflebarger HL, Lenke LG, Lowe TG, Betz RR, Bassett GS. Parents' and patients' preferences and concerns in idiopathic adolescent scoliosis: a cross-sectional preoperative analysis. Spine 2000;25:2392–9.

[12] Carreon LY, Sanders JO, Diab M, Sturm PF, Sucato DJ. Patient satisfaction after surgical correction of adolescent idiopathic scoliosis. Spine 2011;36:965–8.

[13] Lenke LG, Betz RR, Clements D, et al. Curve prevalence of a new classification of operative adolescent idiopathic scoliosis: does classification correlate with treatment? Spine 2002;27:604–11.