Minimum Clinically Important Difference: Current Trends in the Orthopaedic Literature, Part I: Upper Extremity

A Systematic Review

Anne G. Copay, PhD
Andrew S. Chung, DO
Blake Eyberg, MD
Neil Olmscheid, MD
Norman Chutkan, MD
Mark J. Spangehl, MD

Investigation performed at SPIRITT Research, St. Louis, Missouri

Abstract

Background: The minimum clinically important difference (MCID) attempts to define the patient's experience of treatment outcomes. Efforts at calculating the MCID have yielded multiple and inconsistent MCID values. The purposes of this review were to describe the usage of the MCID in the most recent orthopaedic literature, to explain the limitations of its current uses, and to clarify the underpinnings of MCID calculation. Subsequently, we hope that the information presented here will help practitioners to better understand the MCID and to serve as a guide for future efforts to calculate the MCID. The first part of this review focuses on the upper-extremity orthopaedic literature. Part II will focus on the lower-extremity orthopaedic literature.

Methods: A review was conducted of the 2014 to 2016 publications in *The Journal of Arthroplasty, The Journal of Bone & Joint Surgery, The American Journal of Sports Medicine, Foot & Ankle International, Journal of Orthopaedic Trauma, Journal of Pediatric Orthopaedics, and Journal of Shoulder and Elbow Surgery*. Only clinical science articles utilizing patient-reported outcome measure (PROM) scores were included in the analysis. A keyword search was then performed to identify articles that calculated or referenced the MCID. Articles were then further categorized into upper-extremity and lower-extremity publications. MCID utilization in the selected articles was subsequently characterized and recorded.

Results: The MCID was referenced in 129 (7.5%) of 1,709 clinical science articles that utilized PROMs: 52 (40.3%) of 129 were related to the upper extremity, 5 (9.6%) of 52 independently calculated MCID values, and 47 (90.4%) of 52 used previously published MCID values as a gauge of their own results. MCID values were considered or calculated for 16 PROMs; 12 of these were specific to the upper extremity. Six different methods were used to calculate the MCID. Calculated MCIDs had a wide range of values for the same PROM (e.g., 8 to 36 points for Constant-Murley scores and 6.4 to 17 points for American Shoulder and Elbow Surgeons [ASES] scores).

Conclusions: Determining useful MCID values remains elusive and is compounded by the proliferation of PROMs in the field of orthopaedics. The fundamentals of MCID calculation methods should be critically evaluated. If necessary, these methods should be corrected or abandoned.

Disclosure: There was no source of external funding for this study. On the **Disclosure of Potential Conflicts of Interest** forms, *which are provided with the online version of the article*, one or more of the authors checked "yes" to indicate that the author had a relevant financial relationship in the biomedical arena outside the submitted work (http://links.lww.com/JBJSREV/A377).

Furthermore, the type of change intended to be measured should be clarified: beneficial, detrimental, or small or large changes. There should also be assurance that the calculation method actually measures the intended change. Finally, the measurement error should consistently be reported.

Clinical Relevance: The MCID is increasingly used as a measure of patients' improvement. However, the MCID does not yet adequately capture the clinical importance of patients' improvement.

As health-care reimbursement becomes increasingly dependent on the quality of value-based care, patient engagement plays an increasingly larger role in the evaluation of treatment outcomes in the United States[1]. Consequently, the utilization of patient-reported outcome measures (PROMs) has become commonplace. The advent of PROMs has been an attempt to fill the void that exists from the inability of more objective measures to quantify patients' experiences of treatment outcomes. Despite these useful metrics, there are several shortcomings with regard to the usage of PROMs that are worth noting. First, accumulating evidence demonstrates that a wide range of reported improvements may result using the same PROM to assess the same treatment even within similar patient populations. Second, many studies in the literature have shown mean score improvements of whole patient samples, which fundamentally fail to capture the individual patient's experience. Third, many PROM scores lack intuitive clinical relevance; for example, one cannot definitively say whether a pain reduction of 2 of 10 will improve a patient's quality of life. Although efforts to better define the clinical importance of PROM scores have been ongoing, currently, to our knowledge, none have yet yielded consistent and valid results.

The minimum clinically important difference (MCID) is the product of one attempt to define patients' experiences. The term MCID was coined in 1989 by Jaeschke et al. as "the smallest difference in score in the domain of interest which patients perceive as beneficial."[2] In other words, the MCID is a calculated minimum threshold value in an outcome of interest that patients perceive as clinically meaningful. Although both improvement and clinical decline can be defined in this way, the calculation of the former has been the primary focus in the literature. The definition of the MCID appears universally accepted, but, unfortunately, there has been extreme variability in both its measurement and applications.

Two broad categories of calculating the MCID include anchor-based methods and distribution-based methods. Furthermore, there are many subcategories of methods within these 2 general groups. Confusingly, this has yielded a countless number of ways that the MCID can be calculated, each commonly resulting in a different value of the MCID.

In the most commonly used anchor-based approach of calculating the MCID, patient-reported outcomes are paired or anchored to another subjective scale, usually some form of a global rating scale. Changes in scores on the patient-reported outcome are compared with a change in a specified global rating scale, and then various statistical tools are applied to determine the MCID. However, the use of a subjective anchor and the inconsistent manner in which global rating scales are formulated and utilized or relative changes in patient-reported outcome scores are clustered raise concerns with regard to the validity of anchor-based MCIDs.

Distribution-based methods consider that some measurement of variability in PROM scores, such as the standard deviation or the effect size, is indicative of the MCID. Two commonly implemented distribution-based methods of calculating the MCID are the standard error of measurement (SEM) and the minimum (or smallest) detectable change (MDC). The SEM represents the underlying chance for error in a measured score due to the test itself (i.e., due to random variation). The MDC is a variation of the SEM that estimates the smallest amount of change in a score that is not due to chance. A frequently cited flaw of the distribution-based methods is that, because they are primarily statistically based, they do not adequately address patient-perceived clinical importance, the underlying tenet of the MCID.

Unfortunately, there is no consensus on the optimal method of calculating the MCID. Consequently, calculated values of the MCID in the literature vary quite drastically for the same patient-reported outcomes. This variability in MCID values has been shown to be related to a multitude of factors, including the patient population being studied, treatment modality, follow-up intervals, and method used to calculate the MCID. The wide spectrum of MCID calculation methods and its implications for the resultant MCID values have previously been reviewed[3,4].

The consequent existence of multiple and inconsistent MCID values has contributed to confusion among practitioners and scientists. Thus, the purpose of this review was to describe the usage of the MCID in the most recent orthopaedic literature, to explain the limitations of its current uses, and to clarify the underpinnings of the MCID calculation. Ultimately, we hope that our review will help practitioners to better understand the MCID. Because PROMs are joint-specific, this review is divided into 2 parts. In this current part, we will first address the upper-extremity literature. In Part II[5], the lower-extremity literature will be discussed. In this part, Part I, we will discuss the multiplicity of the MCID values and the following concepts as they relate to MCID: the direction of measured change and the

measurement error. The lower-extremity discussion will address equation confusion and the fundamentally flawed usages of the MCID for power analyses and as a benchmark for group means.

Material and Methods

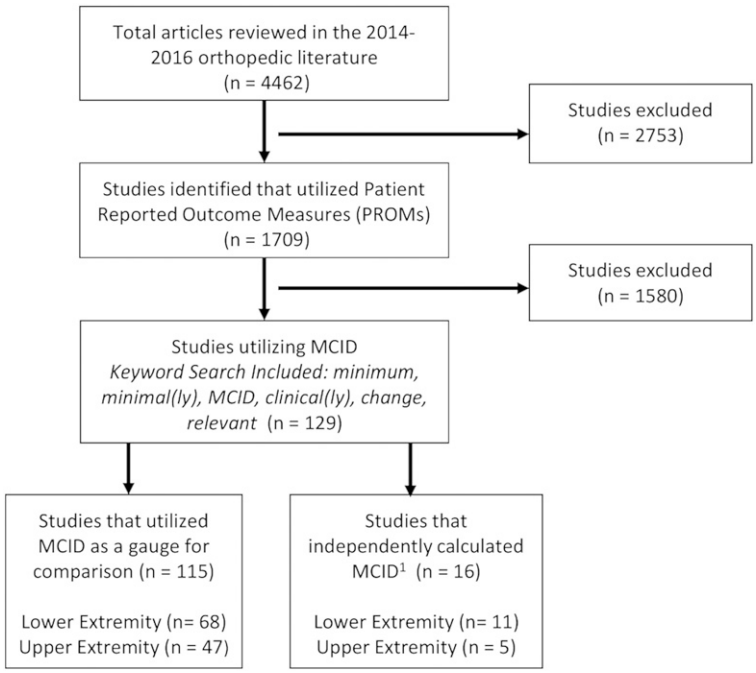
Articles published from 2014 to 2016 in a sampling of 7 major orthopaedic journals were reviewed: the *Journal of Arthroplasty, The Journal of Bone & Joint Surgery, The American Journal of Sports Medicine, Foot & Ankle International, Journal of Orthopaedic Trauma, Journal of Pediatric Orthopaedics,* and *Journal of Shoulder and Elbow Surgery*. We believe that our sampling of general and specialty-specific journals is broadly representative of upper-extremity and lower-extremity subspecialties and accomplishes the task of identifying a random sampling of the orthopaedic literature that utilizes the MCID. Additionally, spine-related publications were not included in either part of this review as they have been previously reviewed in a separate article[4]. Clinical studies utilizing PROMs were identified. A keyword search was performed in each article to select studies that included the MCID. Keywords included "minimum, minimal, minimally, MCID, clinical, clinically, change, relevant" to screen for the most commonly used terms for the MCID. Of these, only studies that mentioned or calculated the MCID were selected. A flow diagram demonstrating the method of study selection is outlined in Figure 1. The same methodology was utilized for both the upper-extremity and lower-extremity portions of our review and, as such, the numbers reported are in relation to all literature reviewed and are subsequently specifically broken down into upper extremity and lower extremity as indicated.

Three types of studies were included in this review: (1) new MCID: 2014 to 2016 studies that calculated new MCID values; (2) MCID gauge: 2014 to 2016 studies that used previously published MCID values as a gauge for their own data; and (3) most cited: prior studies (from any journals and time frames) that calculated the MCID and were most frequently cited by studies in the preceding MCID gauge category [2]).

Studies That Calculated the MCID

The following elements were identified in studies that calculated a new MCID: sample size, patient population, intervention received, length of follow-up, PROM used, MCID calculation method utilized, MCID value, anchor (if used), and measurement error (if calculated).

Anchor-based calculation was the original method utilized to calculate the MCID. In anchor-based MCID calculations, a second subjective patient assessment (known as the anchor or external criterion, also called transition) is used to evaluate the importance of the change seen in PROM scores following an intervention. Most often, the anchor is a global assessment rating scale (see the anchor sections in Table I[6-10] and Table II[11-17]) in which patients may report their outcomes as "much worse" to "much better." The anchor is the basis for the conceptual definition of the MCID, that is, some of the answers on the anchors are believed to represent a minimum important change. Consider the following scenario: A patient reported a change of 3 on a visual analog



[1]Does not include referenced studies that calculated MCID prior to 2014-1016

Fig. 1

Flow diagram illustrating article selection and characteristics of selected articles.

TABLE I 2014 to 2016 Articles Calculating MCID Values for the Upper Extremity*

	Study				
	Wong (2016)[6]	Holmgren (2014)[7]	Castricini (2014)[8]	Torrens (2016)[9]	Somerson (2015)[10]
Sample size	107	93	27	60	17
Diagnosis	Glenohumeral arthritis or cuff tear arthropathy	Subacromial pain	Irreparable rotator cuff tear	Rotator cuff deficiency	Glenohumeral arthritis
Treatment	Shoulder arthroplasty	Exercise	Latissimus dorsi tendon transfer	Shoulder arthroplasty	Glenoid arthroplasty
Mean follow-up	1 yr	3 mo	27 mo	1 yr	3.9 yr
Anchor	Not anchor-based	Global impression of change	Rating of surgery results	Global perception of change −7 to +7[93]	Not anchor-based
Patient subsets (anchor answers)		Importantly improved (recovered, large improvement); not importantly changed (small improvement, unchanged); other (worse)	Very satisfactory, satisfactory, unsatisfactory	No change (hardly better = +1; no change = 0; hardly worse = −1); minimal change (a little better = +3, somewhat better = +2, somewhat worse = −2, a little worse = −3); moderate change (a good deal better = +5, moderately better = +4, moderately worse = −4, a good deal worse = −5); large change (a very great deal better = +7, a great deal better = +6, a great deal worse = −6, a very great deal worse = −7)	
MCID calculation method	0.5 standard deviation	(1) MCID from ROC: importantly improved vs. not importantly changed; (2) MCID = mean change + 1.64 × standard deviation of not importantly changed	MCID = mean change of satisfactory	MCID from ROC: subsets with ≥2 points on the anchor	MCID = 30% of possible maximum improvement
PROM: MCID value	ASES function = 6.5; ASES pain = 8.0; SF-12 PCS = 5.4; SF-12 MCS = 5.7	(1) Constant-Murley = 17 (AUC not reported); Constant-Murley = 22% change from baseline (AUC not reported); (2) Constant-Murley = 24	Constant-Murley = 36	Constant score overall = 8 (AUC not reported)	SST = (maximum SST score − preoperative SST score) × 30%[94]
Measurement error	NR	NR	NR	NR	NR
PROM error value	NR	NR	NR	NR	NR

*NR = not reported.

scale (VAS) for pain following an intervention. Concomitantly, he reported that he was "somewhat better" on a global assessment rating scale. "Somewhat better" is the anchor level that has been predetermined by the clinician or researcher to be indicative of a minimum change. Therefore, in this situation, a VAS of 3 would be considered to be equal to the MCID.

Unfortunately, there is no consensus agreement on what anchor or even specific anchor level (or levels) would best express a minimum important change. Very often, researchers combine several anchor levels as a representation of the minimum important change. As a result, outcomes from "much worse" to "much better" collectively may be considered to express a minimum important change, a clear deviation from the initial utilization of the global assessment rating scale.

Further complicating matters is that even after an anchor level has been selected to represent a minimum change, the MCID may still be calculated in several ways. The following anchor-based methods have been utilized in the literature included in this review. For ease of description, "improved" patients in this discussion are patients who are considered to report a minimum important change, and "non-improved" patients are patients selected as a comparison group to the improved patients.

In the mean change method, the MCID equals the mean change in the PROM score of the improved patients, in other words, the mean score changes of all patients who reported themselves as experiencing a minimum change in the global rating scale following an intervention.

In the change difference method, the MCID equals the difference in the

TABLE II MCID Studies Most Cited by 2014 to 2016 Upper-Extremity Articles*

	Study						
	Tashjian (2010)[11]	Michener (2002)[12]	Kukkonen (2013)[13]	Tashjian (2009)[14]	Roy (2009)[15]	Roy (2010)[16]	Gummesson (2003)[17]
No. of citations	14	9	6	4	5	4	2
Sample size	81	63	781	81	Literature review	120	109
Diagnosis	Rotator cuff pathologies	Shoulder pathologies	Rotator cuff tear	Rotator cuff pathologies		Shoulder pathology	Upper-extremity musculoskeletal conditions
Treatment	Nonoperative	Physical therapy	Arthroscopic tear repair	Nonoperative		Shoulder arthroplasty	Surgery
Follow-up	6 wk	3 to 4 wk	3 mo	6 wk		6 mo	12 mo
Anchor	(1) Global perception of change, −7 to +7; (2) Response to treatment 0-3	Global rating of change	Rating of shoulder	Response to treatment 0 to 3		Change in DASH score	Rating of involved arm
Patient subsets (anchor answers)	(1) No change (hardly better = +1, no change = 0, hardly worse = −1); small change (a little better = +3, somewhat better = +2, somewhat worse = −2, a little worse = −3); (2) no change (none = 0, poor = 1); small change (good = 2)	Improved (much better, slightly better); not improved (same, slightly worse, much worse)	Satisfied (better), dissatisfied (worse)	No change (none= 0, poor =1); small change (good = 2)		≥10.2 or <10.2	Much better, somewhat better, no change, somewhat worse, much worse
MCID calculation method	MCID = change difference between small change and no change	MCID from ROC curve, improved vs. not improved	MCID = mean change of satisfied	MCID = change difference between small change and no change	NR	MCID from ROC curve	MCID = mean change of somewhat better and somewhat worse
PROM: MCID value	SST = 2; ASES = 12 to 17	ASES = 6.4 (AUC = 0.818)	Constant score = 10.4	VAS pain = 1.4 cm (on a 10-cm scale)	ASES = 6.4; DASH = 10.2; SPADI = 8 to 13; SST: no existing value	SST = 3.0 (AUC = 0.661)	DASH = 10
Measurement error	NR	(1) SEM, (2) MDC_{90}	NR	NR	(1) MDC_{90}; (2) MDC_{95}	NR	NR
PROM: error value	NR	(1) ASES = 6.7, (2) ASES = 9.4 (or 9.7; both values were found in article)	NR	NR	(1) ASES = 9.4; DASH = 10.5; (2) SPADI = 18; SST: no existing value	NR	NR

*NR = not reported; DASH = Disabilities of the Arm, Shoulder and Hand; and SPADI = Shoulder Pain and Disability Index.

mean change in the PROM score between the improved and non-improved patients, in other words, the mean score change of all patients who reported themselves as experiencing a minimum change in the global rating scale minus the mean score change of all patients who did not report themselves as experiencing a minimum change.

In the receiver operating characteristic (ROC) curve method, the MCID equals the cutoff score between improved and non-improved patients determined by an ROC curve. The ROC curve is another modality that identifies a change in a PROM that discriminates between the improved and non-improved patients. It relies on the area under the curve (AUC) to determine how often the score change correctly discriminates between 2 patients. For instance, if the AUC equals 0.80, a specific change in score correctly identifies improved and non-improved patients 80% of the time. AUC values approaching 1.0 represent excellent discrimination.

In the mean change limit method, the MCID equals the mean change (as above, but of non-improved patients) + 1.645 times the standard deviation of the non-improved patients' reported score[18].

Distribution-based MCID calculations compare the change in PROM scores with some measure of variability such as the standard deviation or effect size. Distribution-based methods do not conceptually define the minimum change (in contrast to anchor-based calculation methods). The following distribution-based methods were utilized in the literature included in this review.

In the 0.5 standard deviation method, the MCID value is equal to half of the standard deviation of the measured change in PROM scores[19].

In the percentage from baseline method, the MCID equals a specific percentage of improvement from the baseline score, for instance, an improvement of 30% from before the intervention.

The measurement error and the MCID are different concepts; nonetheless,

the calculation of MCID values for a PROM should take measurement error (i.e., the imprecision of that PROM) into account. The SEM has traditionally been utilized as a representation of the random variation in a data set of recorded scores. Hence, a change in a score that is smaller than the calculated SEM is likely to occur because of measurement error and is less likely to represent a true change. Another measure of measurement error, the MDC, is the smallest change that can be considered to exceed the measurement error with a given level of confidence (usually at a 95% level of confidence; sometimes, 90%). Ultimately, if the value calculated for the MCID is smaller than that of the measurement error, the PROM may not be responsive enough to detect an improvement that is deemed meaningful to the individual patient[20].

The SEM and MDC are calculated as follows[21]:

- $SEM = SD\sqrt{1-r}$ (SD: standard deviation of baseline scores, r: test-retest reliability coefficient)
- $MDC_{90} = 1.65\sqrt{2}\,SEM$ (hence, $MDC_{90} = 2.33$ SEM)
- $MDC_{95} = 1.96\sqrt{2}\,SEM$ (hence, $MDC_{95} = 2.77$ SEM)

Tables I and II include the following information: the anchor type (for anchor-based calculations) with the anchor answers; the combination of anchor answers to form subsets of improved and non-improved patients, the MCID calculation method utilized, and the measurement error (if applicable).

Studies That Used Published MCID Values as a Gauge

For studies that used a previously published MCID value as a gauge for their own data, the following elements of that study were reported: the orthopaedic subspecialty, the PROM used, the MCID value and the corresponding article(s) referenced, and how the MCID was used in the study. The MCID had 3 possible uses: (1) sample size: studies based their power analysis on previously reported MCID values to determine the sample size required to detect a change in score or a difference between groups equal to the MCID value; (2) group means: studies compared the mean score improvement in their sample to the previously reported MCID; and (3) individual scores: studies calculated the proportion of patients whose score improvement reached a previously reported MCID.

Studies That Were Most Frequently Referenced

Previously published studies were cited as references by the studies using the MCID as a gauge. The most frequently cited of these studies were analyzed similarly to the studies calculating the MCID and the following elements were identified: sample size, patient population, length of follow-up, PROM used, MCID calculation method utilized, MCID value, potential anchor, and measurement error.

Results

In total, 4,462 articles were reviewed. Of these, 1,709 were clinical science articles that utilized PROMs as part of their study design. A total of 129 (7.5%) of the 1,709 studies used or referenced the MCID to some extent; 52 (40.3%) of 129 were related to the upper extremity (Fig. 1). Five studies calculated new MCID values (Table I), 47 studies used the MCID as a gauge (see Appendix[22-68], with a number of source articles[11-17,19,69-88]), and 7 studies were the most frequently cited in the reviewed literature (Table II). Altogether, 11 studies calculated the MCID (5 studies calculated a new MCID and 6 of the most frequently cited studies also calculated the MCID). These 11 studies used a total of 6 different methods; the ROC curve was the most frequently used method of MCID calculation. Only 2 studies showed estimates of the measurement error; in both cases, the MCID was within the measurement error[12,15]. MCID values were calculated for the Constant-Murley score, Simple Shoulder Test (SST), American Shoulder and Elbow Surgeons (ASES) score, Short Form-12 (SF-12) items, and VAS for pain. Multiple calculations for the same questionnaire yielded a wide range of MCID values, for example, 8 to 36 points for the Constant-Murley score and 6.4 to 17 points for the ASES score. The reported MCID values used as a gauge also varied noticeably. Altogether, MCID values were considered or were calculated for 16 PROMs; 12 of these were joint-specific PROMs for the upper extremity.

When the MCID was used as a gauge, 55.3% of studies compared the mean score change between preoperative and postoperative PROM scores with the MCID, 40.4% used the MCID to determine their sample size, and 10.6% showed the proportion of patients in their studies who improved to the level of the MCID after the studied intervention. A total of 28 different studies were cited as references for the MCID; of those, only 7 were cited more than once (Table II).

Discussion

The current climate regarding the MCID is limited by disorganization and a general lack of a consensus agreement on its appropriate usage. This can certainly be expected given the multitude of factors that contribute to variability in MCID values and what seems to be a widespread misunderstanding with regard to the fundamentals of MCID usage. Further, the multiplicity of PROMs, heterogeneity of published MCID values, and inconsistency in the calculation of MCID have made interpretation and effective utilization of this potentially powerful metric increasingly difficult. Despite these obvious flaws, there nonetheless appears to be interest in this potentially powerful metric. It can be anticipated that this interest may potentially grow given the current progressive shift toward value-based care in our current health-care economy.

Based on our review, we identified numerous factors that further undermine the utility of the MCID and contribute to the general confusion that underlies MCID utilization. We will elaborate on 3 of these concepts in this part of our 2-part discussion.

The first concept is MCID multiplicity. Several factors are known to contribute to the multiplicity of MCIDs. First, the MCID is always PROM-specific. As such, as more PROMs are developed to evaluate treatment results, the number of potential MCID values will likely increase. Second, for the same PROM, different calculation methods yield variable MCIDs that may have up to a tenfold difference in values[89]. Furthermore, even the same calculation method may result in variable MCID values depending on the characteristics of the sample population, the underlying pathology being addressed, and the treatment being studied[4]. Hence, the MCID should be calculated with a sample that is sufficiently large to minimize the influence of atypical patients. Additionally, when the MCID is used as a gauge, the pathology or treatment under study should be similar to the pathology or treatment that was initially used to calculate the cited MCID value.

The second concept is beneficial compared with detrimental change. In their seminal article, Jaeschke et al.[2] defined the MCID as a beneficial change but proceeded to measure the MCID as any small change, be it improvement or deterioration. A change of 1 to 3 points in either direction was considered a minimal change. The authors only considered the magnitude of the PROM score change, without the direction of the change, so that negative changes did not nullify positive changes in the mean PROM score change. A proportion of MCID studies continue to incorporate worsening as part of the change to be measured but fail to take into account the direction of the change in their calculation of the MCID[9,11,12]. For instance, the most cited MCID value for the ASES score (6.4 points) is the discriminating value between patients who improve and those who get worse[12] and, for that reason, fundamentally is not a measure of beneficial change. Future studies should take care to incorporate only beneficial changes into their calculation of the MCID. Readers should additionally be mindful of this subtle flaw when deciding on whether to utilize a previously published MCID.

The third concept is meaningful compared with random change. Measurement error is typically calculated as the SEM; $MDC_{90} = 2.33$ SEM and $MDC_{95} = 2.77$ SEM. Some of the early MCID research set the MCID equal to 1 SEM[90,91]. Since then, two schools of thought have considered how the MCID and the measurement error should relate. One opinion is that a measurement error, such as the MDC, a statistically based calculation, is fundamentally different from the MCID and, as such, should not be used as a measure of the MCID. The logic for this argument is as follows: although a calculated value of the MCID may be inside the range of measurement error, this value may not necessarily reflect a change that is unimportant to the patient[12,18,92]. Yet another belief is that MCID values should not be lower than the measurement error. This is also our current belief. The rationale for this belief is that it would be impossible to discriminate between random and meaningful change if the MCID value is lower than the measurement error[12].

Given the confusion, most authors continue to disregard the issue altogether. In our review, the measurement error was not reported in 10 of the 12 orthopaedic studies (Tables I and II). Although the relationship between the MCID and the measurement error has not been fully defined[92], it is our recommendation that a measure of the measurement error nonetheless be reported in conjunction with the MCID.

In conclusion, a meaningful interest in the utility of the MCID has been demonstrated in the upper-extremity orthopaedic literature. As mentioned previously, MCID values may vary substantially on the basis of numerous factors, including the underlying patient population being studied, treatment modality, follow-up intervals, and method used to calculate the MCID. Unfortunately, the recent proliferation of both PROMs and MCID calculation methods has created a climate of confusion. Consequently, although the MCID, in theory, is certainly a potentially powerful metric, it seems as although an improved understanding of the fundamentals of the MCID is warranted.

Additionally, numerous methodological errors in the utilization of the MCID need to be addressed. These methods should either be corrected or abandoned. Further, the type of change being measured should be clarified, whether beneficial or detrimental. Finally, until the ideal method of calculating the MCID is identified, we recommend that readers and clinicians utilize the 4 following questions when critically evaluating MCID values. (1) Is the measured change a true change? If the MCID value is less than the MDC_{95}, the change in PROM score is likely a random change. If the MDC_{95} is not reported, it is not possible to assess whether the MCID represents a true change. (2) Is the measured change minimal? If MCID is calculated for patients who report being much better or largely improved, it is unlikely that the MCID represents a minimal change. (3) Is the measured change important? If the MCID is calculated for patients who report no change, it is unlikely that the change is important. Furthermore, if patients who report being worse are included in the calculation, the MCID no longer represents a desirable change. (4) Is the sample used for calculating the MCID large and representative of a specific pathology or treatment?

Furthermore, it has been our recommendation that the MDC_{95} be reported in conjunction with the MCID or utilized as a surrogate for the MCID. This, at the very least, ensures that the calculated MCID value is at least free of measurement error (i.e., it is not simply due to random variation).

Appendix

A table showing 2014 to 2016 upper-extremity articles using the MCID as a gauge is available with the online version of this article as a data supplement at jbjs.org (http://links.lww.com/JBJSREV/A378).

Anne G. Copay, PhD[1],
Andrew S. Chung, DO[2],
Blake Eyberg, MD[3],
Neil Olmscheid, MD[4],
Norman Chutkan, MD[3],
Mark J. Spangehl, MD[2]

[1]SPIRITT Research, St. Louis, Missouri

[2]Department of Orthopaedics, Mayo Clinic-Arizona, Phoenix, Arizona

[3]Orthopaedic Surgery Residency, University of Arizona College of Medicine, Phoenix, Arizona

[4]Orthopedic Surgery Residency, McLaren Greater Lansing, Michigan State University, Lansing, Michigan

E-mail address for A.G. Copay: acopay@spirittresearch.com

ORCID iD for A.G. Copay: 0000-0002-3179-2587
ORCID iD for A.S. Chung: 0000-0003-3942-9595
ORCID iD for B. Eyberg: 0000-0002-9361-1964
ORCID iD for N. Olmscheid: 0000-0002-3970-7850
ORCID iD for N. Chutkan: 0000-0001-7751-9095
ORCID iD for M.J. Spangehl: 0000-0003-1090-9165

References

1. Centers for Medicare & Medicaid Services. CMS quality measure development plan: supporting the transition to the Merit-based Incentive Payment System (MIPS) and Alternative Payment Models (APMs). 2016 May 2. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Value-Based-Programs/MACRA-MIPS-and-APMs/Final-MDP.pdf. Accessed 2018 Feb 16.

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

3. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. Spine J. 2007 Sep-Oct; 7(5):541-6. Epub 2007 Apr 2.

4. Chung AS, Copay AG, Olmscheid N, Campbell D, Walker JB, Chutkan N. Minimum clinically important difference: Current trends in the spine literature. Spine (Phila Pa 1976). 2017 Jul 15;42(14):1096-105.

5. Copay AG, Eyberg B, Chung AS, Zurcher KS, Chutkan N, Spangehl MG. Minimal clinically important difference: current trends in the orthopaedic literature, part II: lower extremity. A systematic review. JBJS Rev. 2018 Sept;6(9):e2.

6. Wong SE, Zhang AL, Berliner JL, Ma CB, Feeley BT. Preoperative patient-reported scores can predict postoperative outcomes after shoulder arthroplasty. J Shoulder Elbow Surg. 2016 Jun; 25(6):913-9. Epub 2016 Mar 31.

7. Holmgren T, Oberg B, Adolfsson L, Björnsson Hallgren H, Johansson K. Minimal important changes in the Constant-Murley score in patients with subacromial pain. J Shoulder Elbow Surg. 2014 Aug;23(8):1083-90. Epub 2014 Apr 13.

8. Castricini R, Longo UG, De Benedetto M, Loppini M, Zini R, Maffulli N, Denaro V. Arthroscopic-assisted latissimus dorsi transfer for the management of irreparable rotator cuff tears: short-term results. J Bone Joint Surg Am. 2014 Jul 16;96(14):e119.

9. Torrens C, Guirro P, Santana F. The minimal clinically important difference for function and strength in patients undergoing reverse shoulder arthroplasty. J Shoulder Elbow Surg. 2016 Feb;25(2):262-8. Epub 2015 Sep 28.

10. Somerson JS, Wirth MA. Self-assessed and radiographic outcomes of humeral head replacement with nonprosthetic glenoid arthroplasty. J Shoulder Elbow Surg. 2015 Jul; 24(7):1041-8. Epub 2015 Jan 1.

11. Tashjian RZ, Deloach J, Green A, Porucznik CA, Powell AP. Minimal clinically important differences in ASES and simple shoulder test scores after nonoperative treatment of rotator cuff disease. J Bone Joint Surg Am. 2010 Feb; 92(2):296-303.

12. Michener LA, McClure PW, Sennett BJ. American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form, patient self-report section: reliability, validity, and responsiveness. J Shoulder Elbow Surg. 2002 Nov-Dec;11(6):587-94.

13. Kukkonen J, Kauko T, Vahlberg T, Joukainen A, Äärimaa V. Investigating minimal clinically important difference for Constant score in patients undergoing rotator cuff surgery. J Shoulder Elbow Surg. 2013 Dec;22(12):1650-5. Epub 2013 Jul 12.

14. Tashjian RZ, Deloach J, Porucznik CA, Powell AP. Minimal clinically important differences (MCID) and patient acceptable symptomatic state (PASS) for visual analog scales (VAS) measuring pain in patients treated for rotator cuff disease. J Shoulder Elbow Surg. 2009 Nov-Dec;18(6):927-32. Epub 2009 Jun 16.

15. Roy JS, MacDermid JC, Woodhouse LJ. Measuring shoulder function: a systematic review of four questionnaires. Arthritis Rheum. 2009 May 15;61(5):623-32.

16. Roy JS, Macdermid JC, Faber KJ, Drosdowech DS, Athwal GS. The simple shoulder test is responsive in assessing change following shoulder arthroplasty. J Orthop Sports Phys Ther. 2010 Jul;40(7):413-21.

17. Gummesson C, Atroshi I, Ekdahl C. The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: longitudinal construct validity and measuring self-rated health change after surgery. BMC Musculoskelet Disord. 2003 Jun 16;4:11. Epub 2003 Jun 16.

18. de Vet HC, Ostelo RW, Terwee CB, van der Roer N, Knol DL, Beckerman H, Boers M, Bouter LM. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. Qual Life Res. 2007 Feb;16(1):131-42. Epub 2006 Oct 11.

19. 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. 2003 May;41(5): 582-92.

20. Hägg O, Fritzell P, Nordwall A; Swedish Lumbar Spine Study Group. The clinical importance of changes in outcome scores after treatment for chronic low back pain. Eur Spine J. 2003 Feb;12(1):12-20. Epub 2002 Oct 24.

21. Beaton DE. Understanding the relevance of measured change through studies of responsiveness. Spine (Phila Pa 1976). 2000 Dec 15;25(24):3192-9.

22. Werner BC, Wong AC, Mahony GT, Craig EV, Dines DM, Warren RF, Gulotta LV. Causes of poor postoperative improvement after reverse total shoulder arthroplasty. J Shoulder Elbow Surg. 2016 Aug;25(8):e217-22. Epub 2016 Apr 7.

23. Shannon SF, Wagner ER, Houdek MT, Cross WW 3rd, Sánchez-Sotelo J. Reverse shoulder arthroplasty for proximal humeral fractures: outcomes comparing primary reverse arthroplasty for fracture versus reverse arthroplasty after failed osteosynthesis. J Shoulder Elbow Surg. 2016 Oct;25(10):1655-60. Epub 2016 Apr 18.

24. Denard PJ, Lädermann A. Immediate versus delayed passive range of motion following total shoulder arthroplasty. J Shoulder Elbow Surg. 2016 Dec;25(12):1918-24. Epub 2016 Oct 7.

25. Chalmers PN, Saltzman BM, Feldheim TF, Mascarenhas R, Mellano C, Cole BJ, Romeo AA, Nicholson GP. A comprehensive analysis of pectoralis major transfer for long thoracic nerve palsy. J Shoulder Elbow Surg. 2015 Jul;24(7): 1028-35. Epub 2015 Feb 2.

26. Li F, Zhu Y, Lu Y, Liu X, Wu G, Jiang C. Hemiarthroplasty for the treatment of complex proximal humeral fractures: does a trabecular metal prosthesis make a difference? A prospective, comparative study with a minimum 3-year follow-up. J Shoulder Elbow Surg. 2014 Oct;23(10):1437-43.

27. Athwal GS, MacDermid JC, Reddy KM, Marsh JP, Faber KJ, Drosdowech D. Does bony increased-offset reverse shoulder arthroplasty decrease scapular notching? J Shoulder Elbow Surg. 2015 Mar;24(3):468-73. Epub 2014 Oct 22.

28. Chan K, MacDermid JC, Hoppe DJ, Ayeni OR, Bhandari M, Foote CJ, Athwal GS. Delayed versus early motion after arthroscopic rotator cuff repair: a meta-analysis. J Shoulder Elbow Surg. 2014 Nov;23(11):1631-9. Epub 2014 Aug 13.

29. Vavken P, Tepolt FA, Kocher MS. Open inferior capsular shift for multidirectional shoulder instability in adolescents with generalized ligamentous hyperlaxity or Ehlers-Danlos syndrome. J Shoulder Elbow Surg. 2016 Jun;25(6):907-12. Epub 2016 Jan 14.

30. Abrams GD, Gupta AK, Hussey KE, Tetteh ES, Karas V, Bach BR Jr, Cole BJ, Romeo AA, Verma NN. Arthroscopic repair of full-thickness rotator cuff tears with and without acromioplasty: randomized prospective trial with 2-year follow-up. Am J Sports Med. 2014 Jun;42(6): 1296-303. Epub 2014 Apr 14.

31. Steen BM, Cabezas AF, Santoni BG, Hussey MM, Cusick MC, Kumar AG, Frankle MA. Outcome and value of reverse shoulder arthroplasty for treatment of glenohumeral osteoarthritis: a matched cohort. J Shoulder Elbow Surg. 2015 Sep;24(9):1433-41. Epub 2015 Mar 11.

32. Kruse LM, Yamaguchi K, Keener JD, Chamberlain AM. Clinical outcomes after decompression of the nerve to the teres minor

in patients with idiopathic isolated teres minor fatty atrophy. J Shoulder Elbow Surg. 2015 Apr; 24(4):628-33. Epub 2014 Oct 29.

33. Pappou I, Virani NA, Clark R, Cottrell BJ, Frankle MA. Outcomes and costs of reverse shoulder arthroplasty in the morbidly obese: a case control study. J Bone Joint Surg Am. 2014 Jul 16;96(14):1169-76.

34. LeBlanc JE, MacDermid JC, Faber KJ, Drosdowech DS, Athwal GS. Outcomes after hemiarthroplasty for proximal humerus fracture are significantly affected by hand dominance. J Orthop Trauma. 2015 Aug;29(8): 379-83.

35. Mulligan EP, Brunette M, Shirley Z, Khazzam M. Sleep quality and nocturnal pain in patients with shoulder disorders. J Shoulder Elbow Surg. 2015 Sep;24(9):1452-7. Epub 2015 Apr 1.

36. Miller RM, Popchak A, Vyas D, Tashman S, Irrgang JJ, Musahl V, Debski RE. Effects of exercise therapy for the treatment of symptomatic full-thickness supraspinatus tears on in vivo glenohumeral kinematics. J Shoulder Elbow Surg. 2016 Apr;25(4):641-9. Epub 2015 Nov 24.

37. Huang H, Grant JA, Miller BS, Mirza FM, Gagnier JJ. A systematic review of the psychometric properties of patient-reported outcome instruments for use in patients with rotator cuff disease. Am J Sports Med. 2015 Oct; 43(10):2572-82. Epub 2015 Jan 26.

38. Keener JD, Galatz LM, Teefey SA, Middleton WD, Steger-May K, Stobbs-Cucchi G, Patton R, Yamaguchi K. A prospective evaluation of survivorship of asymptomatic degenerative rotator cuff tears. J Bone Joint Surg Am. 2015 Jan 21;97(2):89-98.

39. Wylie JD, Suter T, Potter MQ, Granger EK, Tashjian RZ. Mental health has a stronger association with patient-reported shoulder pain and function than tear size in patients with full-thickness rotator cuff tears. J Bone Joint Surg Am. 2016 Feb 17;98(4):251-6.

40. Rubright J, Kelleher P, Beardsley C, Paller D, Shackford S, Beynnon B, Shafritz A. Long-term clinical outcomes, motion, strength, and function after total claviculectomy. J Shoulder Elbow Surg. 2014 Feb;23(2):236-44. Epub 2013 Aug 12.

41. Russell RD, Knight JR, Mulligan E, Khazzam MS. Structural integrity after rotator cuff repair does not correlate with patient function and pain: a meta-analysis. J Bone Joint Surg Am. 2014 Feb 19;96(4):265-71.

42. Moosmayer S, Lund G, Seljom US, Haldorsen B, Svege IC, Hennig T, Pripp AH, Smith HJ. Tendon repair compared with physiotherapy in the treatment of rotator cuff tears: a randomized controlled study in 103 cases with a five-year follow-up. J Bone Joint Surg Am. 2014 Sep 17;96(18):1504-14.

43. Shields E, Mirabelli M, Amsdell S, Thorsness R, Goldblatt J, Maloney M, Voloshin I. Functional and imaging outcomes of arthroscopic simultaneous rotator cuff repair and bankart repair after shoulder dislocations. Am J Sports Med. 2014 Nov;42(11):2614-20. Epub 2014 Sep 26.

44. Hsu JE, Gorbaty JD, Whitney IJ, Matsen FA 3rd. Single-stage revision is effective for failed shoulder arthroplasty with positive cultures for Propionibacterium. J Bone Joint Surg Am. 2016 Dec 21;98(24):2047-51.

45. Hartzler RU, Steen BM, Hussey MM, Cusick MC, Cottrell BJ, Clark RE, Frankle MA. Reverse shoulder arthroplasty for massive rotator cuff tear: risk factors for poor functional improvement. J Shoulder Elbow Surg. 2015 Nov;24(11):1698-706. Epub 2015 Jul 11.

46. Young BL, Watson SL, Meyer RD, Ponce BA. Nonunion of first rib fracture in a softball player: case report of a rare cause of thoracic outlet syndrome. J Shoulder Elbow Surg. 2016 Nov; 25(11):e353-7.

47. van der Meijden OA, Houwert RM, Hulsmans M, Wijdicks FJ, Dijkgraaf MG, Meylaerts SA, Hammacher ER, Verhofstad MH, Verleisdonk EJ. Operative treatment of dislocated midshaft clavicular fractures: plate or intramedullary nail fixation? A randomized controlled trial. J Bone Joint Surg Am. 2015 Apr 15;97(8):613-9.

48. Wellman DS, Lazaro LE, Cymerman RM, Axelrad TW, Leu D, Helfet DL, Lorich DG. Treatment of olecranon fractures with 2.4- and 2.7-mm plating techniques. J Orthop Trauma. 2015 Jan;29(1):36-43.

49. Brehmer JL, Husband JB. Accelerated rehabilitation compared with a standard protocol after distal radial fractures treated with volar open reduction and internal fixation: a prospective, randomized, controlled study. J Bone Joint Surg Am. 2014 Oct 1;96(19):1621-30.

50. Olsen JR, Shields E, Williams RB, Miller R, Maloney M, Voloshin I. A comparison of cortical button with interference screw versus suture anchor techniques for distal biceps brachii tendon repairs. J Shoulder Elbow Surg. 2014 Nov;23(11):1607-11. Epub 2014 Sep 11.

51. Nelson GN, Stepan JG, Osei DA, Calfee RP. The impact of patient activity level on wrist disability after distal radius malunion in older adults. J Orthop Trauma. 2015 Apr;29(4): 195-200.

52. Osei DA, Calfee RP, Stepan JG, Boyer MI, Goldfarb CA, Gelberman RH. Simultaneous bilateral or unilateral carpal tunnel release? A prospective cohort study of early outcomes and limitations. J Bone Joint Surg Am. 2014 Jun 4; 96(11):889-96. Epub 2014 Jun 4.

53. Dunn WR, Kuhn JE, Sanders R, An Q, Baumgarten KM, Bishop JY, Brophy RH, Carey JL, Holloway GB, Jones GL, Ma CB, Marx RG, McCarty EC, Poddar SK, Smith MV, Spencer EE, Vidal AF, Wolf BR, Wright RW. Symptoms of pain do not correlate with rotator cuff tear severity: a cross-sectional study of 393 patients with a symptomatic atraumatic full-thickness rotator cuff tear. J Bone Joint Surg Am. 2014 May 21;96(10):793-800.

54. Okoroha KR, Lynch JR, Keller RA, Korona J, Amato C, Rill B, Kolowich PA, Muh SJ. Liposomal bupivacaine versus interscalene nerve block for pain control after shoulder arthroplasty: a prospective randomized trial. J Shoulder Elbow Surg. 2016 Nov;25(11):1742-8. Epub 2016 Jul 14.

55. Kukkonen J, Joukainen A, Lehtinen J, Mattila KT, Tuominen EK, Kauko T, Äärimaa V. Treatment of nontraumatic rotator cuff tears: a randomized controlled trial with two years of clinical and imaging follow-up. J Bone Joint Surg Am. 2015 Nov 4;97(21):1729-37.

56. Gracitelli ME, Malavolta EA, Assunção JH, Kojima KE, dos Reis PR, Silva JS, Ferreira Neto AA, Hernandez AJ. Locking intramedullary nails compared with locking plates for two- and three-part proximal humeral surgical neck fractures: a randomized controlled trial. J Shoulder Elbow Surg. 2016 May;25(5):695-703.

57. Lambers Heerspink FO, van Raay JJ, Koorevaar RC, van Eerden PJ, Westerbeek RE, van 't Riet E, van den Akker-Scheek I, Diercks RL. Comparing surgical repair with conservative treatment for degenerative rotator cuff tears: a randomized controlled trial. J Shoulder Elbow Surg. 2015 Aug;24(8):1274-81.

58. Russell S, Jariwala A, Conlon R, Selfe J, Richards J, Walton M. A blinded, randomized, controlled trial assessing conservative management strategies for frozen shoulder. J Shoulder Elbow Surg. 2014 Apr;23(4):500-7.

59. Louer CR, Boone SL, Guthrie AK, Motley JR, Calfee RP, Wall LB. Postural stability in older adults with a distal radial fracture. J Bone Joint Surg Am. 2016 Jul 20;98(14):1176-82.

60. Rudge WB, Sewell MD, Al-Hadithy N, Wallace CN, Lambert SM. Shoulder linked arthroplasty in patients with obstetric brachial plexus palsy can improve quality of life and function at short-term follow-up. J Shoulder Elbow Surg. 2015 Sep;24(9): 1473-80. Epub 2015 May 7.

61. Singh HP, Mehta SS, Pandey R. A preoperative scoring system to select patients for arthroscopic subacromial decompression. J Shoulder Elbow Surg. 2014 Sep;23(9):1251-6.

62. Hollman F, Wessel RN, Wolterbeek N. Response shift of the Western Ontario Rotator Cuff index in patients undergoing arthroscopic rotator cuff repair. J Shoulder Elbow Surg. 2016 Dec;25(12):2011-8. Epub 2016 Jul 14.

63. Rasmussen JV, Olsen BS, Al-Hamdani A, Brorson S. Outcome of revision shoulder arthroplasty after resurfacing hemiarthroplasty in patients with glenohumeral osteoarthritis. J Bone Joint Surg Am. 2016 Oct 5;98(19):1631-7.

64. Buckley T, Miller R, Nicandri G, Lewis R, Voloshin I. Analysis of subscapularis integrity and function after lesser tuberosity osteotomy versus subscapularis tenotomy in total shoulder arthroplasty using ultrasound and validated clinical outcome measures. J Shoulder Elbow Surg. 2014 Sep;23(9):1309-17. Epub 2014 Mar 4.

65. Shaha JS, Cook JB, Song DJ, Rowles DJ, Bottoni CR, Shaha SH, Tokish JM. Redefining "critical" bone loss in shoulder instability: functional outcomes worsen with "subcritical" bone loss. Am J Sports Med. 2015 Jul;43(7): 1719-25. Epub 2015 Apr 16.

66. Benegas E, Ferreira Neto AA, Gracitelli ME, Malavolta EA, Assunção JH, Prada FdeS, Bolliger Neto R, Mattar R Jr. Shoulder function after surgical treatment of displaced fractures of the humeral shaft: a randomized trial comparing antegrade intramedullary nailing with minimally invasive plate osteosynthesis. J Shoulder Elbow Surg. 2014 Jun;23(6):767-74. Epub 2014 Apr 24.

67. London DA, Stepan JG, Boyer MI, Calfee RP. The impact of depression and pain catastrophization on initial presentation and treatment outcomes for atraumatic hand conditions. J Bone Joint Surg Am. 2014 May 21;96(10):806-14.

68. Kim HM, Caldwell JM, Buza JA, Fink LA, Ahmad CS, Bigliani LU, Levine WN. Factors affecting satisfaction and shoulder function in patients with a recurrent rotator cuff tear. J Bone Joint Surg Am. 2014 Jan 15;96(2):106-12.

69. Michener LA, Snyder AR, Leggin BG. Responsiveness of the numeric pain rating scale in patients with shoulder pain and the effect of surgical status. J Sport Rehabil. 2011 Feb;20(1):115-28.

70. Kirkley A, Griffin S, Dainty K. Scoring systems for the functional assessment of the shoulder. Arthroscopy. 2003 Dec;19(10):1109-20.

71. Schmitt JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. J Clin Epidemiol. 2004 Oct;57(10):1008-18.

72. MacDermid JC, Drosdowech D, Faber K. Responsiveness of self-report scales in patients recovering from rotator cuff surgery. J Shoulder Elbow Surg. 2006 Jul-Aug;15(4):407-14.

73. Wright RW, Baumgarten KM. Shoulder outcomes measures. J Am Acad Orthop Surg. 2010 Jul;18(7):436-44.

74. Smith MV, Calfee RP, Baumgarten KM, Brophy RH, Wright RW. Upper extremity-specific measures of disability and outcomes in orthopaedic surgery. J Bone Joint Surg Am. 2012 Feb 1;94(3):277-85.

75. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. J Hand Ther. 2001 Apr-Jun;14(2):128-46.

76. Hudak PL, Amadio PC, Bombardier C; The Upper Extremity Collaborative Group (UECG). Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected] [corrected]. Am J Ind Med. 1996 Jun;29(6):602-8.

77. Grewal R, Athwal GS, MacDermid JC, Faber KJ, Drosdowech DS, El-Hawary R, King GJ. Single versus double-incision technique for the repair of acute distal biceps tendon ruptures: a randomized clinical trial. J Bone Joint Surg Am. 2012 Jul 3;94(13):1166-74.

78. Sorensen AA, Howard D, Tan WH, Ketchersid J, Calfee RP. Minimal clinically important differences of 3 patient-rated outcomes instruments. J Hand Surg Am. 2013 Apr; 38(4):641-9. Epub 2013 Mar 6.

79. Kelly AM. The minimum clinically significant difference in visual analogue scale pain score does not differ with severity of pain. Emerg Med J. 2001 May;18(3):205-7.

80. Calfee RP, Adams AA. Clinical research and patient-rated outcome measures in hand surgery. J Hand Surg Am. 2012 Apr;37(4):851-5.

81. Gallagher EJ, Liebman M, Bijur PE. Prospective validation of clinically important changes in pain severity measured on a visual analog scale. Ann Emerg Med. 2001 Dec;38(6): 633-8.

82. Wilson J, Baker P, Rangan A. Is retrospective application of the Oxford Shoulder Score valid? J Shoulder Elbow Surg. 2009 Jul-Aug;18(4): 577-80. Epub 2009 May 7.

83. van Kampen DA, Willems WJ, van Beers LW, Castelein RM, Scholtes VA, Terwee CB. Determination and comparison of the smallest detectable change (SDC) and the minimal important change (MIC) of four-shoulder patient-reported outcome measures (PROMs). J Orthop Surg Res. 2013 Nov 14;8:40.

84. Wessel RN, Wolterbeek N, Fermont AJ, van Mameren H, Sonneveld H, Griffin S, de Bie RA. The conceptually equivalent Dutch version of the Western Ontario Rotator Cuff Index (WORC)©. BMC Musculoskelet Disord. 2013 Dec 21;14:362.

85. Christiansen DH, Frost P, Falla D, Haahr JP, Frich LH, Svendsen SW. Responsiveness and minimal clinically important change: a comparison between 2 shoulder outcome measures. J Orthop Sports Phys Ther. 2015 Aug; 45(8):620-5. Epub 2015 Jun 25.

86. Kirkley A, Griffin S, McLintock H, Ng L. The development and evaluation of a disease-specific quality of life measurement tool for shoulder instability. The Western Ontario Shoulder Instability Index (WOSI). Am J Sports Med. 1998 Nov-Dec;26(6):764-72.

87. Shauver MJ, Chung KC. The minimal clinically important difference of the Michigan hand outcomes questionnaire. J Hand Surg Am. 2009 Mar;34(3):509-14.

88. Jensen MP, Gammaitoni AR, Olaleye DO, Oleka N, Nalamachu SR, Galer BS. The pain quality assessment scale: assessment of pain quality in carpal tunnel syndrome. J Pain. 2006 Nov;7(11):823-32.

89. 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 Nov-Dec;8(6):968-74. Epub 2008 Jan 16.

90. 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. 1999 May;37(5):469-78.

91. 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. 1999 Sep;52(9):861-73.

92. de Vet HC, Terwee CB. The minimal detectable change should not replace the minimal important difference. J Clin Epidemiol. 2010 Jul;63(7):804-5, author reply :806.

93. Juniper EF, Guyatt GH, Willan A, Griffith LE. Determining a minimal important change in a disease-specific Quality of Life Questionnaire. J Clin Epidemiol. 1994 Jan;47(1):81-7.

94. Gilmer BB, Comstock BA, Jette JL, Warme WJ, Jackins SE, Matsen FA. The prognosis for improvement in comfort and function after the ream-and-run arthroplasty for glenohumeral arthritis: an analysis of 176 consecutive cases. J Bone Joint Surg Am. 2012 Jul 18;94(14):e102.

http://dx.doi.org/10.2106/JBJS.RVW.17.00159

Appendix

TABLE E-1 2014 to 2016 Upper-Extremity Articles Using MCID as a Gauge*

Study	Orthopaedic Area	PROM: MCID Value	Source Article	Use
Werner (2016)[22]	Shoulder arthroplasty	ASES: 12	Michener (2002)[12], Tashjian (2010)[11]	Individual scores
Shannon (2016)[23]	Reverse shoulder arthroplasty	ASES: 10	References not reported	Sample size
Denard (2016)[24]	Shoulder arthroplasty	ASES: 6.4	Tashjian (2010)[11]	Sample size
Chalmers (2015)[25]	Long thoracic nerve palsy	ASES: 12 to 17	Tashjian (2010)[11]	Group mean
Li (2014)[26]	Proximal humeral fractures	ASES: 6.4	Michener (2002)[12]	Sample size
Athwal (2015)[27]	Shoulder arthroplasty	ASES: 6.4	Michener (2002)[12]	Sample size
Chan (2014)[28]	Arthroscopic rotator cuff repair	ASES: 6.4	Roy (2009)[15]	Group mean; meta-analysis
Vavken (2016)[29]	Adolescent shoulder instability	ASES: 6.4	Michener (2002)[12]	Group mean
Abrams (2014)[30]	Arthroscopic rotator cuff repair	ASES: 10	Roy (2009)[15] Tashjian (2010)[11]	Sample size
Steen (2015)[31]	Shoulder arthroplasty	ASES: 12; SST: 2	Tashjian (2010)[11]; Roy (2010)[16]	Sample size
Kruse (2015)[32]	Teres minor fatty atrophy	ASES: 12 to 17; SST: 2	Tashjian (2010)[14]	Group mean
Pappou (2014)[33]	Shoulder arthroplasty	ASES: 6.4 to 17; SST: 2 to 3	Michener (2002)[12]; Tashjian (2010)[11]; Roy (2010)[16]	Group mean
LeBlanc (2015)[34]	Humeral fracture	ASES: 9; DASH: 10; SST: 2	Roy (2009)[15]; Tashjian (2010)[11]	Group mean
Mulligan (2015)[35]	Shoulder disorders	ASES: 6.4 (MDC = 9.4); NRS: 2.17	Michener (2002)[12], Michener (2011)[69]	Group mean
Miller (2016)[36]	Supraspinatus tears	ASES: 6.4; DASH: 10.2; WORC: 245.26	Roy (2009)[15]; Kirkley (2003)[70]	Group mean
Huang (2015)[37]	Review of PROMs	ASES: 6.4 (MDC = 9.4); DASH: 10.2 (MDC = 12.2); SPADI: 13.2 (MDC = 18.1); SST: 2.05 (MDC = 3.27)	Michener (2002)[12]; Schmitt (2004)[71]; MacDermid (2006)[72]	Review
Keener (2015)[38]	Rotator cuff tear	ASES: NR; SST: NR	Tashjian (2010)[11]	Group mean
Wylie (2016)[39]	Rotator cuff tear	ASES: NR; SST: NR; VAS pain: NR; VAS function: NR	MCID values and references not reported	Group mean
Rubright (2014)[40]	Claviculectomy	ASES: 6.4; DASH: 10.2	Wright (2010)[73]	Group mean
Russell (2014)[41]	Rotator cuff tear	ASES: 6.4 to 12; VAS pain: 1.4 (0 to 10)	Michener (2002)[12]; Smith (2012)[74]; Tashjian (2010)[11]; Tashjian (2009)[14]	Group mean; meta-analysis
Moosmayer (2014)[42]	Rotator cuff tear	ASES: 6.4, 12 to 17; Constant score: 10.4	Kukkonen (2013)[13]; Michener (2002)[12]; Tashjian (2010)[11]	Group mean
Shields (2014)[43]	Arthroscopic rotator cuff and Bankart repair	ASES: 12; Constant score: 10	Tashjian (2010)[11]; Kukkonen (2013)[13]	Sample size; individual scores
Hsu (2016)[44]	Shoulder arthroplasty	SST: 3	Roy (2010)[16]	Individual scores
Hartzler (2015)[45]	Shoulder	SST: 2 to 3	Roy (2010)[16]; Tashjian	Individual

	arthroplasty		(2010)[11]	scores
Young (2016)[46]	Rib fracture	SST: 17.0 to 19.4; DASH: 15	Tashjian (2010)[11]; Beaton (2001)[75]	Case report
van der Meijden (2015)[47]	Clavicular fracture	DASH: 6	Hudak (1996)[76]	Sample size
Wellman (2015)[48]	Olecranon fractures	DASH: 10	Gummesson (2003)[17]	Group mean
Brehmer (2014)[49]	Radial fracture	DASH: 10.1	Gummesson (2003)[17]	Sample size; group mean
Olsen (2014)[50]	Biceps tendon repair	DASH: 10.2	Grewal (2012)[77]; Roy (2009)[15]	Sample size; group mean
Nelson (2015)[51]	Radial malunion	QuickDASH: 12; VAS pain: NR	Sorensen (2013)[78]; Tashjian (2009)[14]; Kelly (2001)[79]	Sample size; group mean
Osei (2014)[52]	Carpal tunnel	QuickDASH: 12	Calfee (2012)[80]	Sample size
Dunn (2014)[53]	Rotator cuff repair	VAS pain: 1.4 cm	Tashjian (2009)[14]	Group mean
Okoroha (2016)[54]	Shoulder arthroplasty	VAS pain: 13 mm	Gallagher (2001)[81]	Sample size
Kukkonen (2015)[55]	Rotator cuff tears	VAS pain: NR; Constant score: 10	Tashjian (2009)[14]; Kukkonen (2013)[13]	Group mean
Gracitelli (2016)[56]	Humeral surgical neck fractures	Constant-Murley score: 12	Kukkonen (2013)[13]	Sample size
Lambers Heerspink (2015)[57]	Rotator cuff tear	Constant-Murley score: 10.4	Kukkonen (2013)[13]	Group mean
Russell (2014)[58]	Frozen shoulder	Constant score: 15	Authors' clinical practice	Sample size; group mean
Louer (2016)[59]	Radial fracture	DMA: 22	Norman (2003)[19]	Sample size
Rudge (2015)[60]	Shoulder arthroplasty	OSS: 4.5	Wilson (2009)[82]	Group mean
Singh (2014)[61]	Subacromial decompression	OSS: 6, 12 (authors' preference)	Van Kampen (2013)[83]	Group mean
Hollman (2016)[62]	Arthroscopic rotator cuff repair	WORC: 16.7	Wessel (2013)[84]	Individual scores
Rasmussen (2016)[63]	Shoulder arthroplasty	WOOS: 190	Kukkonen (2013)[14]; Christiansen (2015)[85]	Group mean
Buckley (2014)[64]	Shoulder arthroplasty	WOOS: 15	No reference reported	Group mean
Shaha (2015)[65]	Arthroscopic labral repair	WOSI: 200	Kirkley (1998)[86]	Group mean
Benegas (2014)[66]	Humeral fracture	UCLA: 6	No reference reported	Sample size
London (2014)[67]	Atraumatic hand or wrist conditions	MHQ: 6 to 23	Shauver (2009)[87]; Jensen (2006)[88]	Sample size
Kim (2014)[68]	Rotator cuff tear	VAS satisfaction: 2.5	Expert opinion	Sample size

*DASH = Disabilities of the Arm, Shoulder and Hand; DMA = Dynamic Motion Analysis; QuickDASH = abbreviated version of the DASH Outcome Measure; MHQ = Michigan Hand Outcomes Questionnaire; NR = not reported; NRS = Numerical Rating Scale; OSS = Oxford Shoulder Score; SPADI = Shoulder Pain and Disability Index; UCLA = University of California Los Angeles Activity Level Score; WORC = Western Ontario Rotator Cuff Index; WOOS = Western Ontario Osteoarthritis of the Shoulder Index; and WOSI = Western Ontario Shoulder Instability.