Current Reviews in Musculoskeletal Medicine (2021) 14:27–46
https://doi.org/10.1007/s12178-020-09684-2

OUTCOMES RESEARCH IN ORTHOPEDICS (O AYENI, SECTION EDITOR)

Minimal Clinically Important Difference, Substantial Clinical Benefit, and Patient Acceptable Symptom State of Outcome Measures Relating to Shoulder Pathology and Surgery: a Systematic Review

Favian Su[1] · Sachin Allahabadi[1] · Dale N. Bongbong[2] · Brian T. Feeley[1] · Drew A. Lansdown[1]

Accepted: 17 December 2020 / Published online: 12 January 2021

Abstract

Purpose of the Review To provide a comprehensive summary of available literature on the minimal clinically important difference (MCID), substantial clinical benefit (SCB), and patient acceptable symptom state (PASS) of patient-reported outcome measures (PROMs) for various shoulder conditions and outcomes and to identify factors that influence these metrics.

Recent Findings Over the past 2 years, there has been an increasing interest in utilizing MCID, SCB, and PASS as a gauge to evaluate the success of an intervention for shoulder conditions. Efforts at calculating these thresholds have yielded multiple and inconsistent values and are further compounded by the proliferation of different PROMs in the shoulder literature.

Summary The MCID, SCB, and PASS values of shoulder PROMs vary widely with study-specific characteristics, including patient demographics, shoulder pathology, treatment, shoulder instrument, study methodology, and calculation method. The differences in these factors are not inconsequential and could lead to large discrepancies in threshold values. It is crucial that clinicians are mindful of these variables when designing future studies to calculate these metrics or when utilizing previously published values to determine the success of an intervention.

Keywords Minimal clinical important difference · Substantial clinical benefit · Patient acceptable symptom state · Shoulder · Outcomes

Introduction

Integration of patients' perceptions into the effects of treatment has led to an increase in the utilization of patient-reported outcome measures (PROMs) in orthopedic research studies. These metrics allow for injury- or disease-specific evaluations of patient's conditions, factoring in pain, function, and other components. Most commonly, PROMs are compared between groups or over time to determine statistically significant differences or changes; however, statistically significant changes in outcomes may not always equate to clinical significance [1, 2]. Because the meaning of absolute changes in PROMs is not readily obvious, specific thresholds, such as minimal clinically important difference (MCID), substantial clinical benefit (SCB), and patient acceptable symptom state (PASS), have been determined for outcome scores to more clearly convey clinical relevance.

The concept of MCID was first described by Jaeschke et al. and describes the minimum value over which a patient has determined his or her clinical outcome to be beneficial and meaningful [1]. Two common approaches of deriving MCID are distribution-based and anchor-based methods [3]. Distribution-based methods solely rely on the statistical characteristics of the instrument (e.g., standard error of measurement, effect size, standard deviation, or minimum detectable change) and are generally considered less informative because they do not reflect the patient's perspective [4, 5]. In contrast, anchor-based methods specify the patient's perception of improvement

This article is part of the Topical Collection on *Outcomes Research in Orthopedics*

✉ Drew A. Lansdown
Drew.Lansdown@ucsf.edu

[1] Department of Orthopaedic Surgery, University of California, San Francisco, San Francisco, CA, USA

[2] School of Medicine, University of California, San Diego, San Diego, CA, USA

based off of an external criterion or anchor of pain or function. The relation between the anchor and the PROM is analyzed to establish the smallest change in score that best differentiates meaningful change to the patient [6].

Although the MCID is a key threshold that is being utilized with increasing frequency, it represents more of a floor value rather than a goal in terms of defining clinical success [7]. SCB represents the cutoff value for substantial improvement and is differentiated from MCID by identifying patients who responded "much better" rather than "somewhat better" on an external rating of change scale [6]. Another common metric for determining clinical success when utilizing PROMs is PASS, which is the score above which patients consider themselves well. The PASS is determined from the subset of patients who report that their current state of health is satisfactory after taking into account their activities of daily living, level of pain, and functional impairment [8].

Understanding the MCID, SCB, and PASS of shoulder pathologies is key for interpreting outcomes after treatment for various conditions. Although systematic reviews addressing MCID in shoulder PROMs are available, prior reviews limited their analyses to studies utilizing anchor-based methods or failed to perform a credibility assessment [9–11]. Furthermore, there are no reviews to date assessing the SCB or PASS of shoulder instruments. As such, the purpose of this systematic review is to provide a comprehensive summary of available literature on the MCID, SCB, and PASS for various shoulder conditions and outcomes.

Methods

Search Strategy

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines were used in the design of this study (Fig. 1) [12]. A search was conducted using PubMed and Embase databases through May 26, 2020, to identify studies reporting the MCID, SCB, and PASS of all outcome measures pertaining to shoulder conditions. The full search criteria can be found in the Appendix. A total of 930 articles were identified after removal of duplicates. Inclusion criteria were articles pertaining to MCID, SCB, and PASS of outcome measures related to shoulder pathologies (rotator cuff tear, rotator cuff arthropathy, glenohumeral osteoarthritis, shoulder instability, superior labral anterior to posterior (SLAP) tear, biceps tendinitis, subacromial impingement, shoulder pain, acromioclavicular (AC) joint separation, rheumatic shoulder disease, proximal humerus fractures) or nonoperative or surgical treatment of shoulder conditions (physical therapy, rotator cuff repair, total shoulder arthroplasty (TSA), arthroscopic stabilization, etc.). Exclusion criteria included case reports, reviews, nonhuman studies, biomechanical studies, and scientific meeting abstracts or proceedings. Two authors (F.S. and S.A.) independently screened the titles, abstracts, and full texts. Any discrepancies in inclusion/exclusion were carried to the next round of screening to ensure thoroughness. References of each included study were further screened to capture any publications that may have eluded the original search queries. Fifty-five articles were found to be relevant, and an additional 5 articles were identified from article references.

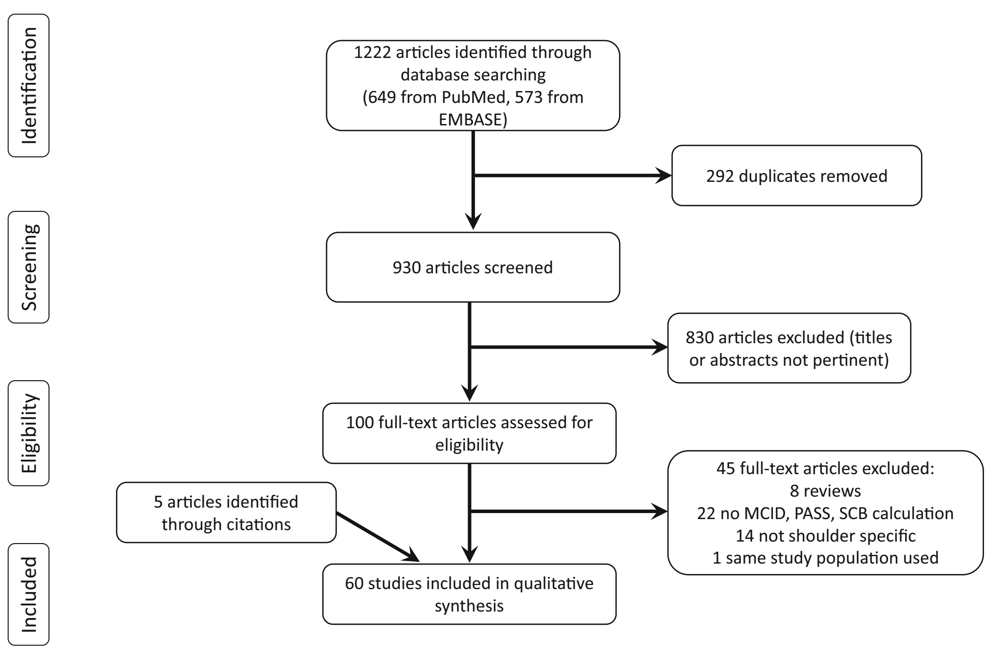

Fig. 1 The Preferred Reporting Items for Systematic Reviews and Meta-Analyses diagram of the literature search and study selection

Data Organization

Relevant data were extracted, including patient demographics, length of follow-up, type of pathology and intervention, anchor information, and calculation method. MCID, SCB, and PASS values were aggregated by PROM instrument, shoulder pathology, and intervention. Mean estimates and ranges were provided for outcome measures.

Credibility Assessment

To assess the extent to which the methodology and performance of studies protect against misleading estimates of MCID and SCB, the credibility of these metrics was evaluated using previously published criteria [13]. A single criterion focusing on the correlation between change in the outcome measure and the anchor (e.g., global rating of change) was used. Values were considered to be credible if the correlation was greater than or equal to 0.4, whereas values were considered to be questionable if the correlation was less than 0.4 or if no correlation was reported [9].

Results

A total of 60 articles were included in this review. Fifty-four (90%) studies reported MCID, nine (15%) calculated SCB, and 11 (18%) quantified PASS. Thirty-two different instruments were utilized; however, only 18 (56%) were reported more than once (Table 1). The sample size and minimum follow-up were highly variable for MCID, SCB, and PASS studies, ranging from 20 to 1568 patients and from 1 week to 24 months, respectively. Nearly all (96%) studies utilized an anchor-based method to calculate MCID; however, there were six different calculations used, including receiver operating characteristic (ROC) (50%), mean difference (24%), mean change (22%), mean change limit (9%), logistic regression (6%), and 75th percentile of the improved group (2%) (Table 2). Most studies used a single anchor that measured global/overall improvement (69%). Seven-point anchors were used most commonly (26%, range, 2–18 points). Distribution-based approaches were utilized less frequently (20%) to determine MCID, with one-half the standard deviation (82%) method being the most common.

For SCB, all studies utilized an anchor-based approach, but only ROC (67%) and mean difference (33%) calculations were used (Table 3). Most studies used a single anchor that measured global/overall improvement with 15-point anchors being the most common (56%, range 4–15 points). Similarly, all PASS studies utilized an anchor-based method assessing satisfaction with only ROC (73%) and 75th percentile of the satisfied group (45%) calculations being used (Table 4). A summary of mean MCID, SCB, and PASS values is shown in Tables 5, 6, 7 and 8. Of the 136 values reported, only 24.6% were considered to be credible.

Eleven studies calculated a MCID for rotator cuff repair with Constant (5 studies), ASES (4 studies), and SANE (3 studies) being the most commonly reported instruments. The mean MCID for Constant, ASES, and SANE using anchor-based methods were 10.9 ± 11.0 (range, 2.0–36.0), 17.8 ± 8.0 (range, 11.1–27.1), and 28.4 ± 1.5 (range, 27.3–29.4), respectively. Conversely, the mean MCID for Constant, ASES, and SANE using distribution-based methods were 6.6 ± 2.8 (range, 4.6–8.6), 19.3 ± 10.7 (range, 11.7–26.9), and 13.5 ± 2.9 (range, 11.8–16.9), respectively. Only one study reported on SCB for rotator cuff repair with mean values of 17.5, 5.5, and 29.8 for ASES, Constant, and SANE. One study reported on PASS for rotator cuff repairs with mean values of 86.7, 23.3, and 82.5 for ASES, Constant, and SANE.

Ten studies calculated a MCID for TSA with ASES (6 studies), SST (5 studies), and Constant (3 studies) being the most commonly reported instruments. The mean MCID for ASES, SST, and Constant using anchor-based methods were 16.0 ± 9.0 (range, 6.3–29.6), 2.9 ± 1.0 (range, 1.5–4.0), and 6.3 ± 1.5 (range, 5.1–8.0), respectively. Conversely, the mean MCID for ASES, SST, and Constant using distribution-based methods were 8.9 ± 2.2 (range, 6.5–11.8), 1.8, and 6.9 ± 3.6 (range, 4.3–9.4), respectively. Three studies reported on SCB for TSA with mean values of 23.9 ± 8.8 (range, 12.0–36.6) and 19.4 ± 0.4 (range, 19.1–19.6) for ASES and Constant, respectively. Three studies reported on PASS for TSA with mean values of 78.6 ± 3.0 (range, 76.0–81.9), 48.8 ± 34.3 (range 24.5–73.0), and 61.8 ± 5.3 (range, 58.0–65.5) for ASES, Constant, and SANE, respectively.

Other operative treatments were reported less frequently. Three studies calculated an MCID for arthroscopic shoulder stabilization. The MCID value for Rowe score was 9.7 and 5.6 for anchor-based and distribution-based calculations, respectively, whereas it was 8.5 for ASES. One study determined the MCID, SCB, and PASS values for biceps tenodesis. The anchor-based MCID values for ASES, Constant, and SANE was 16.3, 6.8, and 3.5, respectively. The SCB threshold was 16.8, 11.0, and 5.8, respectively, whereas the PASS was 59.6, 19.5, and 65.5, respectively. One study estimated the MCID for AC joint stabilization with cutoff for Constant, NPRS, and Taft being 16.6, 1.4, and 2.9, respectively. One study calculated the MCID for SLAP repair using two different methods. The mean MCID for OSIS, Rowe, and WOSI were 9.0 ± 1.4 (range, 8.0–10.0), 17.5 ± 0.7 (range, 17.0–18.0), and 510.0 ± 83.4 (range, 451.0–569.0), respectively.

There were nine studies that determined the MCID for nonoperative management of subacromial impingement and rotator cuff tears with WORC (3 studies), Constant (2 studies), and OSS (2 studies) being the most commonly reported measures. The mean MCID for WORC, Constant, and OSS using anchor-based methods were 343.6 ± 94.3 (range, 269.0–

Table 1 Frequency of instruments reported

Instrument	Abbreviation	No. Studies	Score range	
			Worse shoulder condition	Better shoulder condition
American Shoulder Elbow Surgeons	ASES	18	0	100
Constant-Murley	Constant	16	0	100
Disability of Arm, Shoulder, Hand	DASH	9	100	0
Shoulder Pain and Disability Index	SPADI	9	100	0
Single Alphanumeric Evaluation	SANE	8	0	100
Visual Analog Scale Pain	VAS Pain	8	10	0
Oxford Shoulder Scale	OSS	7	0	48
Simple Shoulder Test	SST	7	0	12
Numeric Pain Rating Scale	NPRS	4	10	0
Disability of Arm, Shoulder, Hand Short Version	QuickDASH	4	100	0
UCLA Shoulder Rating Scale	UCLA	4	2	35
Western Ontario Rotator Cuff Index	WORC	4	2100	0
Rowe Score	Rowe	3	0	100
Western Ontario Shoulder Instability Index	WOSI	3	2100	0
Global Shoulder Function	GSF	2	0	10
Oxford Shoulder Instability Score	OSIS	2	0	48
Patient-Specific Functional Scale	PSFS	2	0	10
Penn Shoulder Score	PSS	2	0	100
Bostrom Shoulder Movement Impairment Scale	Bostrom	1	5	30
Functional Shoulder Scale	FSS	1	0	100
Neer Function Score	Neer	1	100	0
Shoulder Function Index	SFInX	1	0	100
SF-12 Mental Component Score	SF-12 MCS	1	0	100
SF-12 Physical Component Score	SF-12 PCS	1	0	100
Shoulder Disability Questionnaire United Kingdom	SDQ-UK	1	100	0
Shoulder Disability Questionnaire Dutch	SDQ-NL	1	100	0
Shoulder Function Assessment Scale	SFA	1	0	70
Subjective Shoulder Value	SSV	1	0	100
Taft Score	Taft	1	0	12
VR-12 Mental Component Score	VR-12 MCS	1	0	100
VR-12 Physical Component Score	VR-12 PCS	1	0	100
Western Ontario Osteoarthritis Score	WOOS	1	100	0

879.9), 18.5 ± 5.8 (range, 11.0–24.0), and 7.6 ± 3.5 (range, 4.0–12.2), respectively. Two studies reported on SCB for subacromial impingement with values of 11 and 21 for DASH and PSS, respectively. Similarly, two studies reported on PASS for subacromial impingement and rotator cuff tears managed nonoperatively with values of 21.3, 2.3, and 3.0 for Neer function score, NPRS, and VAS pain, respectively.

Nine studies calculated the MCID for physical therapy of nonspecific shoulder pain. The MCID for QuickDASH (3 studies) was 15.7 ± 8.4 (range, 8.0–27.8), whereas for SPADI (3 studies), it was 13.4 ± 6.1 (8.0–20.0). Although no studies on SCB were reported for nonspecific shoulder pain, the SANE and SPADI PASS threshold were 87 and 47.3 ± 1.5 (range, 46.2–48.3). Two studies calculated the MCID of physical therapy for shoulder instability. The MCID of OSIS was 5.3 ± 1.1 (range, 4.5–6.0) and SRQ was 5.0. One study reported on the SCB of physical therapy for shoulder instability with OSIS and SRQ cutoffs of 6.5 and 5.0, respectively. Only one study reported on the MCID of physical therapy for proximal humerus fractures with Constant, DASH, and OSS values of 11.6, 13.0, and 11.4.

Table 2 Characteristics of studies reporting on MCID for shoulder patient-reported outcomes

Study	Year	Instrument(s)	*N*	Follow-up (months)	Diagnosis	Treatment	Method	Domain	Calculation
Rysstad et al. [14]	2020	QuickDASH, PSFS	106	3	Shoulder pain	PT	Anchor	Global: 7 points	(1) ROC (2) Logistic regression
Tashjian et al. [15]	2020	ASES, SST, VAS Pain	202	12	RC tear	RC repair	Anchor	Global: 4 points	Mean difference
Berglund et al. [16]	2019	ASES, SANE, SST, VAS Pain	534	24	OA, RCA	TSA	Anchor	Satisfaction: 4 points	Mean difference
Berglund et al. [17]	2019	SST	176	24	RCA	TSA	Anchor	Satisfaction: 4 points	Mean difference
Cvetanovich et al. [6]	2019	ASES, Constant, SANE	288	12	RC tear	RC repair	Anchor, Distribution	Pain: 15 points	(1) ROC (2) ½ × SD
Gowd et al. [18]	2019	ASES, Constant, SANE	207	12	OA, RCA	TSA	Anchor, distribution	Pain: 15 points	(1) ROC (2) Distribution: not specified
Park et al. [19]	2019	ASES, Rowe	216	12	Shoulder instability	Arthroscopic stabilization	Anchor	Function: 15 points	Mean difference
Policastro et al. [20]	2019	ASES	130	–	Various shoulder pathologies	PT	Anchor	Global: 15 points	Mean difference
Puzzitiello et al. [21]	2019	ASES, Constant, SANE	123	6	Biceps tendinitis	Biceps tenodesis	Anchor, distribution	Pain: 15 points	(1) ROC (2) ½ × SD
Kc et al. [22]	2019	SPADI	119	–	Shoulder pain	PT	Anchor	Global: 7 points	ROC
Kc et al. [23]	2019	DASH	121	1	Shoulder pain	PT	Anchor	Global: 7 points	ROC
Xu et al. [24]	2019	Constant, OSS, UCLA	327	24	RC tear	RC repair	Anchor	Fulfillment: 7 points; Satisfaction: 6 points	Logistic regression
Braun et al. [25]	2018	WORC	54	3	RC tear	PT	Anchor	Global: 7 points	Logistic regression
Budtz et al. [26]	2018	QuickDASH	261	3	Shoulder pain	PT	Anchor	Global: 7 points	(1) ROC (2) Mean change +1.645 × SD
Gagnier et al. [27]	2018	ASES, WORC	222	16	RC tear	Operative, nonoperative	Anchor, distribution	Global: 5 points	(1) Mean change (2) ½ × SD
Gowd et al. [28]	2018	ASES, Constant	89	12	RC tear	RC repair	Anchor	Pain: 15 points	ROC
Koorevaar et al. [29]	2018	DASH	144	12	Various shoulder pathologies	Operative	Anchor	Global: 7 points	Mean change
Park et al. [30]	2018	Rowe, WOSI	198	12	Shoulder instability	Arthroscopic stabilization	Anchor	Function: 15 points	Mean difference
Simovitch et al. [31]	2018	ASES, Constant, GSF, SST, SPADI, UCLA, VAS pain	466	24	OA, RCA	TSA	Anchor, distribution	Global: 4 points	(1) Mean difference (2) ½ × SD
Stein et al. [32]	2018	Constant, NPRS, Taft	73	24	AC joint separation	AC stabilization	Anchor	Global: 6 points	Mean difference
Thigpen et al. [33]	2018	SANE	212	3	Various shoulder pathologies	Operative, nonoperative	Distribution	N/A	Distribution: not specified

Table 2 (continued)

Study	Year	Instrument(s)	*N*	Follow-up (months)	Diagnosis	Treatment	Method	Domain	Calculation
Wessel et al. [34]	2018	WORC	105	6	Various shoulder pathologies	Operative, nonoperative	Anchor	Hindrance: 11 points	(1) ROC (2) Mean change +1.645 × SD
Zhou et al. [35]	2018	SANE, VR-12	160	16	RC tear	Operative, nonoperative	Anchor, distribution	Global: 5 points	(1) Mean change (2) ½ × SD
Thoomes-de Graaf et al. [36]	2017	SPADI	237	6.5	Shoulder pain	PT	Anchor	Global: 7 points	ROC
Rysstad et al. [37]	2017	DASH	50	4	Subacromial impingement	PT	Anchor	Global: 3 points	ROC
Tashjian et al. [38]	2017	ASES, SST, VAS pain	326	24	OA, RCA	TSA	Anchor	Global: 4 points	Mean difference
van der Linde et al. [39]	2017	OSS, WOSI	105	6	Shoulder instability	PT	Anchor	Global: 7 points	ROC
Torrens et al. [40]	2016	Constant	60	12	RC tear, RCA	TSA	Anchor	Function: 15 points	ROC
van der Water [41]	2016	SFInX	74	1.75	Proximal humerus fractures	PT	Anchor, distribution	Global: 15 points	(1) Mean change (2) ½ × SD
Werner et al. [42]	2016	ASES	490	24	OA, RCA	TSA	Anchor	Global, work, activities: 5 points	Mean difference
Wong et al. [43]	2016	ASES, SF-12	107	12	OA, RCA	TSA	Distribution	NA	½ × SD
Christiansen et al. [44]	2015	Constant, OSS	126	3	Subacromial impingement	PT	Anchor	Global: 7 points	(1) ROC (2) Mean change +1.645 × SD
Iossifidis et al. [45]	2015	FSS	119	6	RC tear	RC repair	Anchor	Global: 5 points	Mean change
Negahban et al. [46]	2015	SPADI, DASH	200	1	Various shoulder pathologies	PT	Anchor	Global: 7 points	ROC
Castricini et al. [47]	2014	Constant	27	24	RC tear	RC repair	Anchor	Satisfaction: 3 points	Mean change
Holmgren et al. [48]	2014	Constant	93	3	Subacromial impingement, RC tear	PT	Anchor	Global: 5 points	(1) ROC (2) Mean change +1.645 × SD
Koehorst et al. [49]	2014	PSFS	50	1.5	Shoulder pain	PT	Anchor	Global: 11 points	ROC
Lundquist et al. [50]	2014	DASH	59	8	Various shoulder pathologies	PT	Anchor	Global: 7 points	ROC
Skare et al. [51]	2014	OSIS, Rowe, WOSI	89	6	SLAP tear	Operative, nonoperative	Anchor	Global: 4 points	(1) ROC (2) Mean change +1.645 × SD
van der Water [52]	2014	Constant, DASH, OSS, SSV, UCLA	20	3	Proximal humerus fractures	PT	Anchor, Distribution	Global: 15 points	(1) Mean change (2) ½ × SD
Kukkonen et al. [53]	2013	Constant	776	3	RC tear	RC repair	Anchor, Distribution	Global: 2 points	(1) Mean change (2) Mean difference (3) ROC (4) ½ × SD

Table 2 (continued)

Study	Year	Instrument(s)	*N*	Follow-up (months)	Diagnosis	Treatment	Method	Domain	Calculation
van Kampen et al. [54]	2013	DASH, OSS, QuickDASH, SST	164	6	Various shoulder pathologies	Operative, nonoperative	Anchor	Function, pain: 7 points	Mean change
Christie et al. [55]	2011	Bostrom, Constant, DASH, OSS, SFA, SPADI, VAS pain	100	12	Rheumatic shoulder disease	Operative	Anchor	Global: 5 points	Mean change
Michener et al. [56]	2011	NPRS	136	1	Shoulder pain	PT	Anchor	PSS	ROC
Ekeberg et al. [57]	2010	OSS, SPADI, WORC	121	6	Subacromial impingement	Steroid injection	Anchor	Global: 18 points	(1) ROC (2) Mean change (3) Mean change +1.645 × SD
Roy et al. [58]	2010	SST	120	6	OA, RCA	TSA	Anchor	DASH	ROC
Tashjian et al. [59]	2010	ASES, SST	81	1.5	RC tear	PT	Anchor	Global: 4 points; Function, pain: 15 points	Mean difference
Mintken et al. [60]	2009	NPRS, QuickDASH	101	1	Shoulder pain	PT	Anchor	Global: 15 points	ROC
Tashjian et al. [61]	2009	VAS Pain	81	6	Subacromial impingement, RC tear	PT	Anchor	Global: 4 points	Mean difference
Moser et al. [62]	2008	OSIS, SRQ	100	9	Shoulder instability	PT	Anchor	Global: 5 points	ROC
Leggin et al. [63]	2006	PSS	49	1	Various shoulder pathologies	PT	Anchor	Global: 5 points	Mean change
Tubach et al. [8]	2006	Neer, NPRS	252	0.25	Subacromial impingement	NSAIDs	Anchor	Global: 15 points	75th percentile
Paul et al. [64]	2004	SDQ-UK, SDQ-NL, SPADI, SRQ	180	6	Shoulder pain	Nonoperative	Anchor	Global: 7 points	ROC
Michener et al. [65]	2002	ASES	59	1	Various shoulder pathologies	PT	Anchor	Global: 5 points	ROC

MCID, minimal clinically important difference; QuickDASH, Disability of Arm, Shoulder, Hand short version; PSFS, Patient-Specific Functional Scale; PT, physical therapy; ROC, receiver operating characteristic; ASES, American Shoulder and Elbow Surgeons; SST, Simple Shoulder Test; VAS, Visual Analog Scale; RC, rotator cuff; SANE, Single Alpha Numeric Evaluation; OA, osteoarthritis; RCA, rotator cuff arthropathy; TSA, total shoulder arthroplasty; SD, standard deviation; SPADI, Shoulder Pain and Disability Index; DASH, Disability of Arm, Shoulder, Hand; OSS, Oxford Shoulder Score; UCLA, UCLA Shoulder Rating Scale; WORC, Western Ontario Rotator Cuff Index; WOSI, Western Ontario Shoulder Instability Index; GSF, Global Shoulder Function; NPRS, Numeric Pain Rating Scale; VR-12, Veterans Rand 12; SFInX, Shoulder Function Index; SF-12, Short Form-12; FSS, Functional Shoulder Score; OSIS, Oxford Shoulder Instability Score; SSV, Subjective Shoulder Value; SFA, Shoulder Function Assessment Scale; SRQ, Shoulder Rating Questionnaire; PSS, Penn Shoulder Scale; SDQ-UK, Shoulder Disability Questionnaire United Kingdom; SDQ-NL, Shoulder Disability Questionnaire Dutch

Table 3 Characteristics of studies reporting on SCB for shoulder patient-reported outcomes

Study	Year	Instrument(s)	N	Follow-up (months)	Diagnosis	Treatment	Calculation	Domain	Method
Cvetanovich et al. [6]	2019	ASES, Constant, SANE	288	12	RC tear	RC repair	Anchor	Pain: 15 points	ROC
Gowd et al. [18]	2019	ASES, Constant, SANE	207	12	OA, RCA	TSA	Anchor	Pain: 15 points	ROC
Policastro et al. [20]	2019	ASES	130	–	Various shoulder pathologies	PT	Anchor	Global: 15 points	Mean difference
Puzzitiello et al. [21]	2019	ASES, Constant, SANE	123	6	Biceps tendinitis	Biceps tenodesis	Anchor	Pain: 15 points	ROC
Gowd et al. [28]	2018	ASES, Constant	89	12	RC tear	RC repair	Anchor	Pain: 15 points	ROC
Simovitch et al. [66]	2018	ASES, Constant, GSF, UCLA, SST, SPADI, VAS Pain	1568	24	OA, RCA	TSA	Anchor	Global: 4 points	Mean difference
Werner et al. [42]	2016	ASES	490	24	OA, RCA	TSA	Anchor	Global, work, activities: 5 points	Mean difference
Michener et al. [67]	2013	DASH, PSS	74	3	Subacromial impingement	PT	Anchor	Global: 13 points	ROC
Moser et al. [62]	2008	OSIS, SRQ	100	9	Shoulder instability	PT	Anchor	Global: 5 points	ROC

SCB, Substantial Clinical Benefit; ASES, American Shoulder and Elbow Surgeons; SANE, Single Alpha Numeric Evaluation; RC, rotator cuff; ROC, receiver operating characteristic; OA, osteoarthritis; RCA, rotator cuff arthropathy; TSA, total shoulder arthroplasty; PT, physical therapy; GSF, Global Shoulder Function; UCLA, UCLA Shoulder Rating Scale; SST, Simple Shoulder Test; SPADI, Shoulder Pain and Disability Index; VAS, Visual Analog Scale; DASH, Disability of Arm, Shoulder, Hand; PSS, Penn Shoulder Scale; OSIS, Oxford Shoulder Instability Score; SRQ, Shoulder Rating Questionnaire

Table 4 Characteristics of studies reporting on PASS for shoulder patient-reported outcomes

Study	Year	Instrument(s)	N	Follow-up (months)	Diagnosis	Treatment	Calculation	Method
Cvetanovich et al. [6]	2019	ASES, Constant, SANE	288	12	RC tear	RC repair	Anchor	ROC
Gowd et al. [18]	2019	ASES, Constant, SANE	207	12	OA, RCA	TSA	Anchor	ROC
Puzzitiello et al. [21]	2019	ASES, Constant, SANE	123	6	Biceps tendinitis	Biceps tenodesis	Anchor	ROC
Tran et al. [68]	2019	SPADI	304	6	Shoulder pain	Operative, nonoperative	Anchor	(1) ROC (2) 75th percentile
Gowd et al. [28]	2018	ASES, Constant	89	12	RC tear	RC repair	Anchor	ROC
Chamberlain et al. [69]	2017	ASES, SST, VAS Pain	326	24	OA, RCA	TSA	Anchor	75th percentile
Sciascia et al. [70]	2017	ASES, Constant, SANE, WOOS	234	24	OA	TSA	Anchor	ROC
O'Halloran et al. [71]	2013	SANE	68	NR	Shoulder pain	PT	Anchor	ROC
Christie et al. [55]	2011	Bostrom, Constant, DASH, OSS, SFAS, SPADI, VAS pain	100	12	Rheumatic shoulder disease	Operative	Anchor	(1) ROC (2) 75th percentile
Tashjian et al. [61]	2009	VAS Pain	81	6	Subacromial impingement, RC tear	PT	Anchor	75th percentile
Tubach et al. [8]	2006	Neer, NPRS	252	0.25	Subacromial impingement	NSAID	Anchor	75th percentile

PASS, patient acceptable symptom state; ASES, American Shoulder and Elbow Surgeons; SANE, Single Alpha Numeric Evaluation; RC, rotator cuff; ROC, receiver operating characteristic; OA, osteoarthritis; RCA, rotator cuff arthropathy; TSA, total shoulder arthroplasty; SPADI, Shoulder Pain and Disability Index; SST, Simple Shoulder Test; VAS, Visual Analog Scale; WOOS, Western Ontario Osteoarthritis of the Shoulder; PT, physical therapy; DASH, Disability of Arm, Shoulder, Hand; OSS, Oxford Shoulder Score; SFAS, Shoulder Function Assessment Scale; NPRS; Numeric Pain Rating Scale; NSAID, nonsteroidal anti-inflammatory drug

Table 5 MCID values of shoulder assessment instruments for operative shoulder pathologies

	Anchor MCID[a]	Anchor Credibility	Distribution MCID[a]
Rotator cuff repair			
Constant ($n = 5$)	10.9 ± 11.0 (2.0–36.0)		6.6 ± 2.8 (4.6–8.6)
Castricini et al., 2014 [47]	36.0	Questionable	–
Cvetanovich etl al, 2019 [6]	5.5	Questionable	4.6
Gowd et al., 2018 [28]	4.6	Questionable	–
Kukkonen et al., 2013 [53]	Mean change: 10.4 Mean difference: 16.4 ROC: 2	Questionable	8.6
Xu et al., 2019 [24]	Satisfaction: 6.3 Fulfillment: 6.1	Questionable	–
ASES ($n = 4$)	17.8 ± 8.0 (11.1–27.1)		19.3 ± 10.7 (11.7–26.9)
Cvetanovich et al., 2019 [6]	11.1	Questionable	11.7
Gagnier et al., 2018 [27]	21.9	Questionable	26.9
Gowd et al., 2018 [28]	11.1	Questionable	–
Tashjian et al., 2020 [15]	27.1	Questionable	–
SANE ($n = 3$)	28.4 ± 1.5 (27.3–29.4)		13.5 ± 2.9 (11.8–16.9)
Cvetanovich et al., 2019 [6]	29.4	Questionable	16.9
Thigpen et al., 2018 [33]	–	–	11.8
Zhou et al., 2018 [35]	27.3	Questionable	11.8
WORC ($n = 2$)	579.3 ± 257.3 (282.6–741.3)		588.7
Gagnier et al., 2018 [27]	282.6	Questionable	588.7
Wessel et al., 2018 [34]	ROC: 714 Mean change limit: 741.3	Credible	–
FSS ($n = 1$)	24.7		–
Iossifidis et al., 2015 [45]	24.7	Questionable	–
OSS ($n = 1$)	2.7 ± 0.1 (2.6–2.7)		–
Xu et al., 2019 [24]	Satisfaction: 2.6 Fulfillment: 2.7	Questionable	–
SST ($n = 1$)	4.3		–
Tashjian et al., 2020 [15]	4.3	Questionable	–
UCLA ($n = 1$)	2.8 ± 0.1 (2.7–2.9)		–
Xu et al., 2019 [24]	Satisfaction: 2.9 Fulfillment: 2.7	Questionable	–
VAS Pain ($n = 1$)	2.4		–
Tashjian et al., 2020 [15]	2.4	Questionable	–
VR-12 MCS ($n = 1$)	1.9		6
Zhou et al., 2018 [35]	1.9	Questionable	6
VR-12 PCS ($n = 1$)	2.6		4.9
Zhou et al., 2018 [35]	2.6	Questionable	4.9
Total shoulder arthroplasty			
ASES ($n = 6$)	16.0 ± 9.0 (6.3 ± 29.6)		8.9 ± 2.2 (6.5–11.8)
Berglund et al., 2019 [16]	29.6	Questionable	–
Gowd et al., 2019 [18]	27.6	Questionable	9.1
Simovitch et al., 2018 [31]	13.6	Questionable	11.8
Tashjian et al., 2017 [38]	20.9	Questionable	–
Werner et al., 2016 [42]	Work: 6.3 Activities: 9.1 Overall: 13.5 SF-12: 7.7	Questionable	–
Wong et al., 2016 [43]	–	–	Function: 6.5 Pain: 8
SST ($n = 5$)	2.9 ± 1.0 (1.5–4.0)		1.8
Berglund et al., 2019 [16]	3.6	Questionable	–
Berglund et al., 2019 [17]	4	Questionable	–
Roy et al., 2010 [58]	3	Credible	–
Simovitch et al., 2018 [31]	1.5	Questionable	1.8
Tashjian et al., 2017 [38]	2.4	Questionable	–

Table 5 (continued)

	Anchor MCID[a]	Anchor Credibility	Distribution MCID[a]
Constant (*n* = 3)	6.3 ± 1.5 (5.1–8.0)		6.9 ± 3.6 (4.3–9.4)
Gowd et al., 2019 [18]	5.1	Questionable	4.3
Simovitch et al., 2018 [31]	5.7	Questionable	9.4
Torrens et al., 2016 [40]	8	Questionable	–
SANE (*n* = 3)	33.0 ± 5.9 (28.8–37.1)		15.8 ± 3.3 (13.4–18.1)
Berglund et al., 2019 [16]	37.1	Questionable	–
Gowd et al., 2019 [18]	28.8	Questionable	13.4
Thigpen et al., 2018 [33]	–	–	18.1
VAS Pain (*n* = 3)	2.1 ± 1.0 (1.4–3.3)		1.6
Berglund et al., 2019 [16]	3.3	Questionable	–
Simovitch et al., 2018 [31]	1.6	Questionable	1.6
Tashjian et al., 2017 [38]	1.4	Questionable	–
GSF (*n* = 1)	1.4		1.4
Simovitch et al., 2018 [31]	1.4	Questionable	1.4
UCLA (*n* = 1)	8.7		3.6
Simovitch et al., 2018 [31]	8.7	Questionable	3.6
SF-12 MCS (*n* = 1)	–		5.7
Wong et al., 2016 [43]	–	–	5.7
SF-12 PCS (*n* = 1)	–		5.4
Wong et al., 2016 [43]	–	–	5.4
SPADI (*n* = 1)	20.6		14.1
Simovitch et al., 2018 [31]	20.6	Questionable	14.1
Arthroscopic shoulder stabilization			
Rowe (*n* = 2)	9.7		5.6
Park et al., 2019 [19]	9.7	Questionable	–
Park et al., 2018 [30]	–	–	5.6
ASES (*n* = 1)	8.5		–
Park et al., 2019 [19]	8.5	Questionable	–
WORC (*n* = 1)	816.9 ± 210.9 (667.8–966.0)		
Wessel et al., 2018 [34]	ROC: 667.8 Mean change limit: 966	Credible	–
WOSI (*n* = 1)	–		151.9
Park et al., 2018 [30]	–	Questionable	151.9
Various shoulder pathologies			
DASH (*n* = 2)	12.7 ± 0.4 (12.4–13.0)		–
Koorevaar et al., 2018 [29]	13	Questionable	–
van Kampen et al., 2013 [54]	12.4	Questionable	–
OSS (*n* = 1)	5.4 ± 0.9 (4.7–6.0)		–
van Kampen et al., 2013 [54]	Function: 6 Pain: 4.7	Questionable	–
QuickDASH (*n* = 1)	13.4		–
van Kampen et al., 2013 [54]	13.4	Questionable	–
SST (*n* = 1)	2.2		–
van Kampen et al., 2013 [54]	2.2	Questionable	–
Biceps tenodesis			
ASES (*n* = 1)	16.3		11
Puzzitiello et al., 2019 [21]	16.3	Questionable	11
Constant (*n* = 1)	6.8		3.8
Puzzitiello et al., 2019 [21]	6.8	Questionable	3.8
SANE (*n* = 1)	3.5		15.2
Puzzitiello et al., 2019 [21]	3.5	Questionable	15.2
AC joint stabilization			
Constant (*n* = 1)	16.6		–
Stein et al., 2018 [32]	16.6	Questionable	–
NPRS (*n* = 1)	1.4		–

Table 5 (continued)

	Anchor MCID[a]	Anchor Credibility	Distribution MCID[a]
Stein et al., 2018 [32]	1.4	Questionable	–
Taft ($n = 1$)	2.9		–
Stein et al., 2018 [32]	2.9	Questionable	–
SLAP repair			
OSIS ($n = 1$)	9.0 ± 1.4 (8.0–10.0)		
Skare et al., 2014 [51]	ROC: 10 Mean change limit: 8	Credible	–
Rowe ($n = 1$)	17.5 ± 0.7 (17.0–18.0)		
Skare et al., 2014 [51]	ROC: 17 Mean change limit: 18	Credible	–
WOSI ($n = 1$)	510.0 ± 83.4 (451.0–569.0)		
Skare et al., 2014 [51]	ROC: 569 Mean change limit: 451	Credible	–
Rheumatic disease			
Bostrom ($n = 1$)	2.3		–
Christie et al., 2011 [55]	2.3	Questionable	–
Constant ($n = 1$)	16.6		–
Christie et al., 2011 [55]	16.6	Questionable	–
DASH ($n = 1$)	10.1		–
Christie et al., 2011 [55]	10.1	Questionable	–
OSS ($n = 1$)	6.9		–
Christie et al., 2011 [55]	6.9	Questionable	–
SFA ($n = 1$)	12.9		–
Christie et al., 2011 [55]	12.9	Questionable	–
SPADI ($n = 1$)	21.3		–
Christie et al., 2011 [55]	21.3	Questionable	–

[a] Data expressed as mean ± SD (range)

MCID, minimal clinically important difference; ROC, receiver operating characteristic; ASES, American Shoulder and Elbow Surgeons; SANE, Single Alpha Numeric Evaluation; WORC, Western Ontario Rotator Cuff; FSS, Functional Shoulder Scale; OSS, Oxford Shoulder Scale; SST, Simple Shoulder Test; UCLA, UCLA Shoulder Rating Scale; VAS, Visual Analog Scale; VR-12 MCS, Veterans Rand-12 Mental Component Score; VR-12 PCS, Veterans Rand-12 Physical Component Score; GSF, Global Shoulder Function; SF-12 MCS, Short Form-12 Mental Component Score; SF-12 PCS, Short Form-12 Physical Component Score; SPADI, Shoulder Pain and Disability Index; WOSI, Western Ontario Shoulder Instability; DASH, Disability of the Arm, Shoulder, Hand; QuickDASH, Disability of the Arm, Shoulder, Hand short version; NPRS, Numeric Pain Rating Scale; OSIS, Oxford Shoulder Instability Score; SFA, Shoulder Function Assessment

Discussion

The MCID, SCB, and PASS allow for interpretation of PROMs and are important to understand when treating various shoulder conditions. Our review demonstrates that the MCID, SCB, and PASS values vary widely with study-specific characteristics, including patient demographics, shoulder pathology, treatment, shoulder instrument, study methodology, and calculation method. Furthermore, in our appraisal of the literature, approximately 75% of the MCID, SCB, and PASS values were found to have questionable credibility or were not credible due to inadequate reporting. These differences have made interpretation of these metrics increasingly difficult and may potentially undermine the results of studies that utilize these thresholds as a basis for measuring a successful outcome.

One factor contributing to the variability in MCID, SCB, and PASS for the same outcome instrument is the shoulder pathology and treatment. For instance, arthroscopic rotator cuff repairs had larger MCID thresholds for ASES, SANE, WORC, SST, and VAS pain compared to physical therapy for rotator cuff tears. This finding is not unexpected given that patient expectations are likely higher when the intervention is more expensive and riskier [48, 72]. Additionally, patients undergoing rotator cuff repair also required greater improvements compared with patients undergoing TSA in order for their improvement to be considered clinically meaningful. These differences may be explained by the generally younger patient population undergoing rotator cuff repair compared to

Table 6 MCID values of shoulder assessment instruments for nonoperative shoulder pathologies

	Anchor MCID[a]	Anchor credibility	Distribution MCID[a]
Subacromial impingement/rotator cuff tear			
WORC (*n* = 3)	450.8 ± 253.4 (269.0–879.9)		–
Braun et al., 2018 [25]	300	Questionable	–
Ekeberg et al., 2010 [57]	ROC: 269.0 Mean change limit: 324.3	Credible	–
Wessel et al., 2018 [34]	ROC: 480.9 Mean change limit: 879.9	Credible	–
Constant (*n* = 2)	18.5 ± 5.8 (11.0–24.0)		–
Christiansen et al., 2015 [44]	ROC: 11 Mean change limit: 22.1	Credible	–
Holmgren et al., 2014 [48]	ROC: 17 Mean change limit: 24	Credible	–
OSS (*n* = 2)	7.6 ± 3.5 (4.0–12.2)		–
Christiansen et al., 2015 [44]	ROC: 6 Mean change limit: 12.2	Credible	–
Ekeberg et al., 2010 [57]	ROC: 4 Mean change limit: 8.1	Credible	–
ASES (*n* = 1)	15.2 ± 2.8 (12.0–16.9)		
Tashjian et al., 2010 [59]	Function: 12.0 Pain: 16.9 Overall: 16.7	Questionable	–
DASH (*n* = 1)	4.4		–
Rysstad et al., 2017 [37]	4.4	Credible	–
Neer (*n* = 1)	15.1		–
Tubach et al., 2006 [8]	15.1	Questionable	–
NPRS (*n* = 1)	3.4		–
Tubach et al., 2006 [8]	3.4	Questionable	–
SANE (*n* = 1)	–		14
Thigpen et al., 2018 [33]	–	–	14
SPADI (*n* = 1)	22.3 ± 3.3 (20.0–24.6)		–
Ekeberg et al., 2010 [57]	ROC: 20 Mean change limit: 24.6	Credible	–
SST (*n* = 1)	2.2 ± 0.1 (2.1–2.3)		–
Tashjian et al., 2010 [59]	Function: 2.1 Overall: 2.3	Questionable	–
VAS pain	1.4		–
Tashjian et al., 2009 [61]	1.4	Questionable	–
Shoulder pain			
QuickDASH (*n* = 3)	15.7 ± 8.4 (8.0–27.8)		–
Budtz et al., 2018 [26]	ROC: 13.6 Mean change limit: 27.8	Questionable	–
Mintken et al., 2009 [60]	8	Credible	–
Rystaad et al., 2020 [14]	13.6	Credible	–
SPADI (*n* = 3)	13.4 ± 6.1 (8.0–20.0)		–
Thoomes-de Graaf et al., 2017 [36]	20	Credible	–
Paul et al., 2004 [64]	8	Credible	–
Kc et al., 2019 [22]	12.3	Credible	–
NPRS (*n* = 2)	1.7 ± 0.8 (1.1–2.2)		–
Michener et al., 2011 [56]	2.2	Questionable	–
Mintken et al., 2009 [60]	1.1	Questionable	–
PSFS (*n* = 2)	1.7 ± 0.5 (1.3–2.0)		–
Koehorst et al., 2014 [49]	1.3	Credible	–
Rysstad et al., 2020 [14]	2	Credible	–
DASH (*n* = 1)	11.2		–
Kc et al., 2019 [23]	11.2	Credible	–
SDQ-NL (*n* = 1)	14		–
Paul et al., 2004 [64]	14	Credible	–
SDQ-UK (*n* = 1)	4		–
Paul et al., 2004 [64]	4	Credible	–
SRQ (*n* = 1)	13		–
Paul et al., 2004 [64]	13	Credible	–

Table 6 (continued)

	Anchor MCID[a]	Anchor credibility	Distribution MCID[a]
Various shoulder pathologies			
ASES (*n* = 2)	7.2 ± 1.1 (6.4–7.9)		–
Michener et al., 2002 [65]	6.4	Credible	–
Policastro et al., 2019 [20]	7.9	Credible	–
DASH (*n* = 2)	18.6 ± 9.7 (11.7–25.4)		–
Lundquist et al., 2014 [50]	11.7	Credible	–
Negahban et al., 2015 [46]	25.4	Credible	–
PSS (*n* = 1)	11.4		–
Leggin et al., 2006 [63]	11.4	Credible	–
SPADI (*n* = 1)	14.9		–
Negahban et al., 2015 [46]	14.9	Credible	–
Shoulder instability			
OSIS (*n* = 2)	5.3 ± 1.1 (4.5–6.0)		–
Moser et al., 2008 [62]	4.5	Questionable	–
van der Linde et al., 2017 [39]	6.0	Credible	–
SRQ (*n* = 1)	5.0		–
Moser et al., 2008 [62]	5.0	Questionable	–
WOSI (*n* = 1)	294.0		–
van der Linde et al., 2017 [39]	294.0	Credible	–
Proximal humerus fractures			
Constant (*n* = 1)	11.6		5.4
van der Water et al., 2014 [52]	11.6	Questionable	5.4
DASH (*n* = 1)	13.0		8.1
van der Water et al., 2014 [52]	13.0	Questionable	8.1
OSS (*n* = 1)	11.4		5.1
van der Water et al., 2014 [52]	11.4	Questionable	5.1
SFInX (*n* = 1)	10.3		11.7
van der Water et al., 2016 [41]	10.3	Questionable	11.7
SSV (*n* = 1)	26.6		12.1
van der Water et al., 2014 [52]	26.6	Questionable	12.1
UCLA (*n* = 1)	2.4		2.0
van der Water et al., 2014 [52]	2.4	Questionable	2.0
Adhesive capsulitis			
SANE (*n* = 1)	–		17.5
Thigpen et al., 2018 [33]	–	–	17.5

[a] Data expressed as mean ± SD (range)

MCID, minimal clinically important difference; WORC, Western Ontario Rotator Cuff; OSS, Oxford Shoulder Score; ASES, American Shoulder and Elbow Surgeons; DASH, Disability of Arm, Shoulder, and Hand; NPRS, Numeric Pain Rating Scale; SANE, Single Alpha Numeric Evaluation; SPADI, Shoulder Pain and Disability Index; SST, Single Shoulder Test; VAS, Visual Analog Scale; QuickDASH, Disability of Arm, Shoulder, and Hand short version; PSFS, Patient-Specific Functional Scale; SDQ-NL, Shoulder Disability Questionnaire Dutch; SDQ-UK, Shoulder Disability United Kingdom; SRQ, Shoulder Rating Questionnaire; PSS, Penn Shoulder Scale; OSIS, Oxford Shoulder Instability Score; WOSI, Western Ontario Shoulder Instability; SFInX, Shoulder Function Index; SSV, Subjective Shoulder Value; UCLA, UCLA Shoulder Rating Scale

TSA and the differing expectations in pain and function after shoulder surgery between the two groups [15]. Furthermore, smaller differences may be clinically significant when symptoms are more severe, as evidenced by lower preoperative outcome scores in TSA patients compared to those undergoing rotator cuff repair [15, 38, 48]. In the present review, nine (15%) studies had also grouped patients with different shoulder conditions and treatments, including operative and nonoperative, together [20, 29, 33, 34, 46, 50, 54, 63, 65]. The MCID values derived from these studies may not be applicable to studies focused on a specific condition or treatment.

The heterogeneity of calculation methods used to derive MCID, SCB, and PASS also contributes to the wide range of values observed. Six different anchor-based approaches were used in 96% of studies, whereas only one distribution-based approach was used in 20% of studies. The decreased popularity of distribution-based methods may be due to the fact that they are generally considered less informative than anchor-based estimates because they rely on the statistical properties of a distribution rather than the patient's perception of improvement [73]. Beaton et al. also showed in a cohort of patients with shoulder pain undergoing physical therapy that the thresholds for defining

Table 7 SCB values for shoulder instruments

	SCB[a]	Anchor credibility
Rotator cuff repair		
ASES ($n = 1$)	17.5	
Cvetanovich et al., 2019 [6]	17.5	Questionable
Constant ($n = 1$)	5.5	
Cvetanovich et al., 2019 [6]	5.5	Questionable
SANE ($n = 1$)	29.8	
Cvetanovich et al., 2019 [6]	29.8	Questionable
Total shoulder arthroplasty		
ASES ($n = 3$)	23.8 ± 8.8 (12.0–36.6)	
Gowd et al., 2019 [18]	20.7	Questionable
Simovitch et al., 2018 [66]	31.5	Questionable
Werner et al., 2016 [42]	Work: 22.3 Activities: 20.2 Overall: 36.6 SF-12: 12.0	Questionable
Constant ($n = 2$)	19.4 ± 0.4 (19.1–19.6)	
Gowd et al., 2019 [18]	19.6	Questionable
Simovitch et al., 2018 [66]	19.1	Questionable
GSF ($n = 1$)	3.1	
Simovitch et al., 2018 [66]	3.1	Questionable
SANE ($n = 1$)	50.2	
Gowd et al., 2019 [18]	50.2	Questionable
SPADI ($n = 1$)	45.4	
Simovitch et al., 2018 [66]	45.4	Questionable
SST ($n = 1$)	3.4	
Simovitch et al., 2018 [66]	3.4	Questionable
UCLA ($n = 1$)	12.6	
Simovitch et al., 2018 [66]	12.6	Questionable
VAS Pain ($n = 1$)	3.2	
Simovitch et al., 2018 [66]	3.2	Questionable
Biceps tenodesis		
ASES ($n = 1$)	16.8	
Puzzitiello et al., 2019 [21]	16.8	Questionable
Constant ($n = 1$)	11	
Puzzitiello et al., 2019 [21]	11	Questionable
SANE ($n = 1$)	5.8	
Puzzitiello et al., 2019 [21]	5.8	Questionable
Subacromial impingement (nonoperative)		
DASH ($n = 1$)	11	
Michener et al., 2013 [67]	11	Questionable
PSS ($n = 1$)	21	
Michener et al., 2013 [67]	21	Questionable
Shoulder instability (nonoperative)		
OSIS ($n = 1$)	6.5	
Moser et al., 2008 [62]	6.5	Questionable
SRQ ($n = 1$)	5	
Moser et al., 2008 [62]	5	Questionable
Various shoulder pathologies (nonoperative)		
ASES ($n = 1$)	9.2	
Policastro et al., 2019 [20]	9.2	Credible

[a] Data expressed as mean ± SD (range)

SCB, Substantial Clinical Benefit; ASES, American Shoulder and Elbow Surgeons; SANE, Single Alpha Numeric Evaluation; GSF, Global Shoulder Function; SPADI, Shoulder Pain and Disability Index; SST, Simple Shoulder Test; UCLA, UCLA Shoulder Rating Scale; VAS, Visual Analog Scale; DASH, Disability of Arm, Shoulder, and Hand; PSS, Penn Shoulder Scale; OSIS, Oxford Shoulder Instability Scale; SRQ, Shoulder Rating Questionnaire

Table 8 PASS values for shoulder instruments

	PASS[a]
Rotator cuff repair	
ASES (*n* = 2)	86.7
Cvetanovich et al., 2019 [6]	86.7
Gowd et al., 2018 [28]	86.7
Constant (*n* = 2)	23.3
Cvetanovich et al., 2019 [6]	23.3
Gowd et al., 2018 [28]	23.3
SANE (*n* = 1)	82.5
Cvetanovich et al., 2019 [6]	82.5
Total shoulder arthroplasty	
ASES (*n* = 3)	78.6 ± 3.0 (76.0–81.9)
Chamberlain et al., 2017 [69]	76.0
Gowd et al., 2019 [18]	81.9
Sciascia et al., 2017 [70]	78.0
Constant (*n* = 2)	48.8 ± 34.3 (24.5–73.0)
Gowd et al., 2019 [18]	24.5
Sciascia et al., 2017 [70]	73.0
SANE (*n* = 2)	61.8 ± 5.3 (58.0–65.5)
Gowd et al., 2019 [18]	65.5
Sciascia et al., 2017 [70]	58.0
SST (*n* = 1)	8.4
Chamberlain et al., 2017 [69]	8.4
VAS Pain (*n* = 1)	1.5
Chamberlain et al., 2017 [69]	1.5
WOOS (*n* = 1)	18.0
Sciascia et al., 2017 [70]	18.0
Biceps tenodesis	
ASES (*n* = 1)	59.6
Puzzitiello et al., 2019 [21]	59.6
Constant (*n* = 1)	19.5
Puzzitiello et al., 2019 [21]	19.5
SANE (*n* = 1)	65.5
Puzzitiello et al., 2019 [21]	65.5
Rheumatic disease	
Bostrom (*n* = 1)	19.3 ± 1.1 (18.5–20.0)
Christie et al., 2011 [55]	ROC: 18.5 75th Percentile: 20
Constant (*n* = 1)	43.0 ± 1.4 (42.0–44.0)
Christie et al., 2011 [55]	ROC: 44.0 75th Percentile: 42
DASH (*n* = 1)	43.0 ± 0.1 (42.9–43.0)
Christie et al., 2011 [55]	ROC: 43.0 75th Percentile: 42.9
OSS (*n* = 1)	27.0 ± 1.3 (26.0–27.9)
Christie et al., 2011 [55]	ROC: 27.9 75th Percentile: 26.0
SFA (*n* = 1)	49.3 ± 4.3 (46.2–52.3)
Christie et al., 2011 [55]	ROC: 46.2 75th Percentile: 52.3
SPADI (*n* = 1)	37.4 ± 5.2 (33.7–41.0)
Christie et al., 2011 [55]	ROC: 41.0 75th Percentile: 33.7
Subacomial impingement/rotator cuff tear (nonoperative)	
Neer (*n* = 1)	21.3
Tubach et al., 2006 [8]	21.3
NPRS (*n* = 1)	2.3
Tubach et al., 2006 [8]	2.3
VAS Pain (*n* = 1)	3
Tashjian et al., 2009 [61]	3.0
Shoulder pain (nonoperative)	
SANE (*n* = 1)	87.0
O'Halloran et al., 2013 [71]	87.0
SPADI (*n* = 1)	47.3 ± 1.5 (46.2–48.3)
Tran et al., 2019 [68]	ROC: 48.3 75th Percentile: 46.2

[a] Data expressed as mean ± SD (range)

PASS, patient acceptable symptom state; ASES, American Shoulder and Elbow Surgeons; SANE, Single Alpha Numeric Evaluation; SST, Simple Shoulder Test; VAS, Visual Analog Scale; WOOS, Western Ontario Osteoarthritis of the Shoulder; DASH, Disability of Arm, Shoulder, Hand; OSS, Oxford Shoulder Score; SFA, Shoulder Function Assessment; SPADI, Shoulder Pain Disability Index; NPRS, Numeric Pain Rating Scale

an important response to treatment differed depending on the technique used [74]. Furthermore, these differences were not inconsequential and could have profound effects on the interpretation of responder-type analysis. These findings were corroborated by Kukkonen et al. who demonstrated an eightfold difference in MCID of Constant score between mean difference method and ROC analysis in 781 patients undergoing arthroscopic rotator cuff repair [53]. Similarly, multiple studies that calculated MCID values utilizing both ROC analysis and the mean change limit method for a variety of shoulder pathologies also showed differences up to 54% [26, 34, 44, 48, 51, 57]. All calculations methods also run the risk of classifying patients as meaningfully improved when the changes in their scores fall within the measurement errors of the data set. To this end, measurement errors, such as the minimal detectable change (MDC), should be reported in conjunction with the MCID, SCB, or PASS values to help differentiate meaningful from random change [75].

In addition to the calculation method, MCID, SCB, or PASS values vary based on the type of anchor used, the definition used for the anchor, and the patient groups being studied [73]. Prior studies have demonstrated that domain-specific questions have higher construct validity as anchors for determining clinically important differences than global transition questions [76]. Despite this, 69% of the studies utilized global ratings of change anchors that focused on overall improvement. Ideally, for questionnaires like ASES which consists of pain and function domains, the external anchor should ask about changes in pain and function. Another source of variation among studies is how MCID, SCB, and PASS are defined. Multiple studies defined unchanged and minimally improved groups differently with 20 (37%) studies incorporating any improvement beyond unchanged (e.g., “completely recovered” or “very great deal better”) into minimally improved groups [14, 22, 23, 25, 26, 34, 36, 37, 39, 44, 46, 48–51, 53, 60, 62, 64]. This may incorrectly increase MCID values, as those levels of improvement reflect substantial clinical benefit rather than minimal clinical improvement. Conversely, eight (15%) studies incorporated patients who did worse (e.g., “much worse” or “very great deal worse”) into the unchanged group [14, 25, 46, 49, 53, 62, 64, 65]. Although the number of patients who did worse after treatment is less, this classification may falsely lower the mean score of the unchanged group and thereby incorrectly increase MCID and SCB values. Furthermore, six (11%) studies measured MCID as any small change, be it improvement or deterioration, compared to the unchanged group [19, 30, 40, 41, 52, 59]. This does not represent a measure of beneficial change, and future studies should be mindful of this discrepancy when deciding to utilize a previously published MCID value.

Although both SCB and PASS have been described for nearly two decades, their reporting in literature has severely lagged behind that of MCID [7, 8, 62, 77]. Among studies with the same cohort, SCB values were approximately 1.6-fold greater than MCID values for arthroscopic rotator cuff repair and 1.7- to 2.7-fold greater than MCID values for TSA [6, 18, 31, 42, 66]. Simovitch et al. reported that a mean improvement of 30% of the total metric value would likely achieve SCB in seven outcome metrics among TSA patients [66]. Similarly, PASS values ranged from 67 to 87% of the total metric value for arthroscopic rotator cuff repair and 58 to 82% of the total metric value for TSA [6, 18, 69, 70]. Future studies calculating the MCID should include SCB and PASS estimates as they provide a spectrum of clinically meaningful outcomes that may be used to counsel patients regarding expectations after shoulder surgery.

While there are multiple studies evaluating the MCID, SCB, and PASS of rotator cuff repair and TSA, few studies have calculated the threshold values for other common shoulder procedures, such as arthroscopic shoulder stabilization, biceps tenodesis, AC joint stabilization, SLAP repair, and open reduction internal fixation of proximal humerus fractures. Currently, most studies on these treatments are limited to small case series performed at a single institution [19, 21, 32]. More studies with larger sample sizes need to be dedicated to these shoulder conditions and treatments, and thereafter, clinically meaningful values can be better established.

Thirty-two PROMs were identified in this review with ASES, Constant, and DASH being the most commonly reported instruments. Interestingly, 44% of the measures only had one study reporting MCID, SCB, or PASS values, suggesting that a large proportion of these outcome measures have not been adopted by the scientific community. The American Shoulder and Elbow Surgeons Value Committee has recently recommended the use of eight shoulder outcome instruments (ASES, OSS, SANE, VR-12, WORC, WOSI, WOOS, and PSS) based on freedom from clinician input, standardization, ease of use, and validation [78]. The Constant score, which was among the most frequently reported instruments in this review, was not recommended due to requiring clinician input to measure strength and motion and the poor standardization and precision of these measurements [79]. As such, continued efforts to utilize the recommended measures in patients with shoulder conditions may potentially limit the number of unnecessary MCID, SCB, and PASS values in future studies.

Despite being a comprehensive review of MCID, SCB, and PASS for shoulder outcome measures, this study is not without limitations. First, the included studies were so heterogeneous that the results varied widely and were difficult to integrate. Mean estimates and ranges for MCID, SCB, and PASS were presented, but future studies need to be cautious prior to utilizing a certain value. Researchers must consider the multitude of factors that affect these metrics, including patient characteristics, study size, pathology, intervention, length of follow-up, and calculation methods [80]. Second, the overwhelming majority of studies failed to report measurement errors, such as MDC, which makes it difficult to determine whether these thresholds are meaningful changes or simply due to random variation. Additionally, the criteria chosen to evaluate the credibility of the MCID, SCB, and PASS thresholds were stringent. Other factors that may contribute to the credibility include whether the anchor addresses the patient’s perspective, the precision of the MCID estimate, and whether the threshold or difference between groups represented a small but important change [10]. Lastly, several studies determined MCID, SCB, and PASS values for different subgroups of their patient populations, such as anatomic versus reverse TSA. Only the overall results are presented here to enable comparison.

In conclusion, the present review provides both anchor-based and distribution-based estimates for MCID, SCB, and PASS of outcome instruments addressing patients with shoulder conditions. ASES, Constant, and DASH were the most frequently utilized instruments, whereas rotator cuff repair and TSA were the most commonly analyzed interventions. There were numerous methodological limitations of the primary studies that resulted in a wide range of values. Additionally, it was observed that there is a paucity of literature that reports the results for SCB and PASS estimates in patients with shoulder disorders. Information from this review is vital for clinicians to appropriately establish patient expectations for recovery.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12178-020-09684-2.

Authors' Contributions Favian Su—Methodology, literature search, data collection, data analysis, writing (original draft preparation).

Sachin Allahabadi—Methodology, literature search, data collection, data analysis, writing (original draft preparation).

Dale Bongbong—Data collection, data analysis.

Brian T Feeley—Conceptualization, writing (reviewing and editing).

Drew A Lansdown—Conceptualization, writing (reviewing and editing).

Data Availability All data was obtained from published articles online.

Compliance with Ethical Standards

Conflict of Interest Favian Su, Sachin Allahabadi, and Dale Bongbong declare that they have no conflict of interest.

Brian T Feeley has the following disclosures: Consultant for Kaliber, Inc.; hospitality and educational payments from Zimmer Biomet, Inc.; American Orthopedic Society for Sports Medicine: Board or committee member; Current Reviews in Musculoskeletal Medicine: Editorial or governing board; Journal of Shoulder and Elbow Surgery: Editorial or governing board; Orthopedic Research Society: Board or committee member.

Drew A Lansdown has the following disclosures: Received fellowship-related research/education funding from Arthrex, Inc. and Smith & Nephew; educational support from Medwest; hospitality payments from Wright Medical; Arthroscopy Association of North America: Board or committee member, American Orthopedic Society for Sports Medicine: Board or committee member.

References

1. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989;10(4):407–15. https://doi.org/10.1016/0197-2456(89)90005-6.
2. Chan LS. Minimal clinically important difference (MCID)–adding meaning to statistical inference. Am J Public Health. 2013;103(11):e24–5. https://doi.org/10.2105/AJPH.2013.301580.
3. King MT. A point of minimal important difference (MID): a critique of terminology and methods. Expert Rev Pharmacoecon Outcomes Res. 2011;11(2):171–84. https://doi.org/10.1586/erp.11.9.
4. Brozek JL, Guyatt GH, Schunemann HJ. How a well-grounded minimal important difference can enhance transparency of labelling claims and improve interpretation of a patient reported outcome measure. Health Qual Life Outcomes. 2006;4:69. https://doi.org/10.1186/1477-7525-4-69.
5. McGlothlin AE, Lewis RJ. Minimal clinically important difference: defining what really matters to patients. JAMA. 2014;312(13):1342–3. https://doi.org/10.1001/jama.2014.13128.
6. Cvetanovich GL, Gowd AK, Liu JN, Nwachukwu BU, Cabarcas BC, Cole BJ, et al. Establishing clinically significant outcome after arthroscopic rotator cuff repair. J Shoulder Elb Surg. 2019;28(5):939–48. https://doi.org/10.1016/j.jse.2018.10.013.
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. https://doi.org/10.2106/JBJS.G.01095.
8. Tubach F, Dougados M, Falissard B, Baron G, Logeart I, Ravaud P. Feeling good rather than feeling better matters more to patients. Arthritis Rheum. 2006;55(4):526–30. https://doi.org/10.1002/art.22110.
9. Jones IA, Togashi R, Heckmann N, Vangsness CT Jr. Minimal clinically important difference (MCID) for patient-reported shoulder outcomes. J Shoulder Elb Surg. 2020;29(7):1484–92. https://doi.org/10.1016/j.jse.2019.12.033.
10. Hao Q, Devji T, Zeraatkar D, Wang Y, Qasim A, Siemieniuk RAC, et al. Minimal important differences for improvement in shoulder condition patient-reported outcomes: a systematic review to inform a BMJ rapid recommendation. BMJ Open. 2019;9(2):e028777. https://doi.org/10.1136/bmjopen-2018-028777.
11. Dabija DI, Jain NB. Minimal clinically important difference of shoulder outcome measures and diagnoses: a systematic review. Am J Phys Med Rehabil. 2019;98(8):671–6. https://doi.org/10.1097/PHM.0000000000001169.
12. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4:1. https://doi.org/10.1186/2046-4053-4-1.
13. Devji T, Guyatt GH, Lytvyn L, Brignardello-Petersen R, Foroutan F, Sadeghirad B, et al. Application of minimal important differences in degenerative knee disease outcomes: a systematic review and case study to inform BMJ rapid recommendations. BMJ Open. 2017;7(5):e015587. https://doi.org/10.1136/bmjopen-2016-015587.
14. Rysstad T, Grotle M, Klokk LP, Tveter AT. Responsiveness and minimal important change of the QuickDASH and PSFS when used among patients with shoulder pain. BMC Musculoskelet Disord. 2020;21(1):328. https://doi.org/10.1186/s12891-020-03289-z.
15. Tashjian RZ, Shin J, Broschinsky K, Yeh CC, Martin B, Chalmers PN, et al. Minimal clinically important differences in the American shoulder and elbow surgeons, simple shoulder test, and visual analog scale pain scores after arthroscopic rotator cuff repair. J Shoulder Elb Surg. 2020;29(7):1406–11. https://doi.org/10.1016/j.jse.2019.11.018.
16. Berglund DD, Law TY, Rosas S, Kurowicki J, Giveans MR, Mijic D, et al. The procedure value index: a new method for quantifying value in shoulder arthroplasty. J Shoulder Elb Surg. 2019;28(2):335–40. https://doi.org/10.1016/j.jse.2018.07.031.
17. Berglund DD, Mijic D, Law TY, Kurowicki J, Rosas S, Levy JC. Value comparison of humeral component press-fit and cemented techniques in reverse shoulder arthroplasty. J Shoulder Elb Surg. 2019;28(3):496–502. https://doi.org/10.1016/j.jse.2018.08.015.
18. Gowd AK, Charles MD, Liu JN, Lalehzarian SP, Cabarcas BC, Manderle BJ, et al. Single assessment numeric evaluation (SANE) is a reliable metric to measure clinically significant improvements following shoulder arthroplasty. J Shoulder Elb Surg. 2019;28(11):2238–46. https://doi.org/10.1016/j.jse.2019.04.041.
19. Park I, Oh MJ, Shin SJ. Minimal clinically important differences and correlating factors for the Rowe score and the American shoulder and elbow surgeons score after arthroscopic stabilization surgery for anterior shoulder instability. Arthroscopy. 2019;35(1):54–9. https://doi.org/10.1016/j.arthro.2018.08.005.
20. Policastro PO, Pierobon A, Perez J, Novoa GA, Calvo Delfino M, Sajfar ME, et al. Cross-cultural adaptation and validation of the argentine "American shoulder and elbow surgeons, patient self-report section" questionnaire. Musculoskelet Sci Pract. 2019;43:37–44. https://doi.org/10.1016/j.msksp.2019.05.010.
21. Puzzitiello RN, Gowd AK, Liu JN, Agarwalla A, Verma NN, Forsythe B. Establishing minimal clinically important difference, substantial clinical benefit, and patient acceptable symptomatic

state after biceps tenodesis. J Shoulder Elb Surg. 2019;28(4):639–47. https://doi.org/10.1016/j.jse.2018.09.025.
22. Kc S, Sharma S, Ginn K, Almadi T, Reed D. Nepali translation, cross-cultural adaptation and measurement properties of the shoulder pain and disability index (SPADI). J Orthop Surg Res. 2019;14(1):284. https://doi.org/10.1186/s13018-019-1285-8.
23. Kc S, Sharma S, Ginn K, Almadi T, Subedi H, Reed D. Cross-cultural adaptation and measurement properties of the Nepali version of the DASH (disability of arm, shoulder and hand) in patients with shoulder pain. Health Qual Life Outcomes. 2019;17(1):51. https://doi.org/10.1186/s12955-019-1105-1.
24. Xu S, Chen JY, Lie HME, Hao Y, Lie DTT. Minimal clinically important difference of Oxford, Constant, and UCLA shoulder score for arthroscopic rotator cuff repair. J Orthop. 2020;19:21–7. https://doi.org/10.1016/j.jor.2019.11.037.
25. Braun C, Handoll HH. Estimating the minimal important difference for the Western Ontario Rotator Cuff Index (WORC) in adults with shoulder pain associated with partial-thickness rotator cuff tears. Musculoskelet Sci Pract. 2018;35:30–3. https://doi.org/10.1016/j.msksp.2018.02.003.
26. Budtz CR, Andersen JH, de Vos Andersen NB, Christiansen DH. Responsiveness and minimal important change for the quick-DASH in patients with shoulder disorders. Health Qual Life Outcomes. 2018;16(1):226. https://doi.org/10.1186/s12955-018-1052-2.
27. Gagnier JJ, Robbins C, Bedi A, Carpenter JE, Miller BS. Establishing minimally important differences for the American shoulder and elbow surgeons score and the Western Ontario Rotator Cuff Index in patients with full-thickness rotator cuff tears. J Shoulder Elb Surg. 2018;27(5):e160–e6. https://doi.org/10.1016/j.jse.2017.10.042.
28. Gowd AK, Cvetanovich GL, Liu JN, Nwachukwu BU, Cabarcas BC, Cole BJ, et al. Preoperative mental health scores and achieving patient acceptable symptom state are predictive of return to work after arthroscopic rotator cuff repair. Orthop J Sports Med. 2019;7(10):2325967119878415. https://doi.org/10.1177/2325967119878415.
29. Koorevaar RCT, Kleinlugtenbelt YV, Landman EBM, van't Riet E, Bulstra SK. Psychological symptoms and the MCID of the DASH score in shoulder surgery. J Orthop Surg Res. 2018;13(1):246. https://doi.org/10.1186/s13018-018-0949-0.
30. Park I, Lee JH, Hyun HS, Lee TK, Shin SJ. Minimal clinically important differences in Rowe and Western Ontario shoulder instability index scores after arthroscopic repair of anterior shoulder instability. J Shoulder Elb Surg. 2018;27(4):579–84. https://doi.org/10.1016/j.jse.2017.10.032.
31. Simovitch R, Flurin PH, Wright T, Zuckerman JD, Roche CP. Quantifying success after total shoulder arthroplasty: the minimal clinically important difference. J Shoulder Elb Surg. 2018;27(2):298–305. https://doi.org/10.1016/j.jse.2017.09.013.
32. Stein T, Muller D, Blank M, Reinig Y, Saier T, Hoffmann R, et al. Stabilization of acute high-grade acromioclavicular joint separation: a prospective assessment of the clavicular hook plate versus the double double-button suture procedure. Am J Sports Med. 2018;46(11):2725–34. https://doi.org/10.1177/0363546518788355.
33. Thigpen CA, Shanley E, Momaya AM, Kissenberth MJ, Tolan SJ, Tokish JM, et al. Validity and responsiveness of the single alpha-numeric evaluation for shoulder patients. Am J Sports Med. 2018;46(14):3480–5. https://doi.org/10.1177/0363546518807924.
34. Wessel RN, Wolterbeek N, Fermont AJM, Lavrijsen L, van Mameren H, de Bie RA. Responsiveness and disease specificity of the Western Ontario rotator cuff index. J Orthop. 2018;15(2):337–42. https://doi.org/10.1016/j.jor.2018.02.012.
35. Zhou L, Natarajan M, Miller BS, Gagnier JJ. Establishing minimal important differences for the VR-12 and SANE scores in patients following treatment of rotator cuff tears. Orthop J Sports Med. 2018;6(7):2325967118782159. https://doi.org/10.1177/2325967118782159.
36. Thoomes-de Graaf M, Scholten-Peeters W, Duijn E, Karel Y, de Vet HC, Koes B, et al. The responsiveness and interpretability of the shoulder pain and disability index. J Orthop Sports Phys Ther. 2017;47(4):278–86. https://doi.org/10.2519/jospt.2017.7079.
37. Rysstad T, Roe Y, Haldorsen B, Svege I, Strand LI. Responsiveness and minimal important change of the Norwegian version of the disabilities of the arm, shoulder and hand questionnaire (DASH) in patients with subacromial pain syndrome. BMC Musculoskelet Disord. 2017;18(1):248. https://doi.org/10.1186/s12891-017-1616-z.
38. Tashjian RZ, Hung M, Keener JD, Bowen RC, McAllister J, Chen W, et al. Determining the minimal clinically important difference for the American shoulder and elbow surgeons score, simple shoulder test, and visual analog scale (VAS) measuring pain after shoulder arthroplasty. J Shoulder Elb Surg. 2017;26(1):144–8. https://doi.org/10.1016/j.jse.2016.06.007.
39. van der Linde JA, van Kampen DA, van Beers L, van Deurzen DFP, Saris DBF, Terwee CB. The responsiveness and minimal important change of the Western Ontario shoulder instability index and Oxford shoulder instability score. J Orthop Sports Phys Ther. 2017;47(6):402–10. https://doi.org/10.2519/jospt.2017.6548.
40. Torrens C, Guirro P, Santana F. The minimal clinically important difference for function and strength in patients undergoing reverse shoulder arthroplasty. J Shoulder Elb Surg. 2016;25(2):262–8. https://doi.org/10.1016/j.jse.2015.07.020.
41. van de Water AT, Davidson M, Shields N, Evans MC, Taylor NF. The shoulder function index (SFInX): evaluation of its measurement properties in people recovering from a proximal humeral fracture. BMC Musculoskelet Disord. 2016;17:295. https://doi.org/10.1186/s12891-016-1138-0.
42. Werner BC, Chang B, Nguyen JT, Dines DM, Gulotta LV. What change in American shoulder and elbow surgeons score represents a clinically important change after shoulder arthroplasty? Clin Orthop Relat Res. 2016;474(12):2672–81. https://doi.org/10.1007/s11999-016-4968-z.
43. Wong SE, Zhang AL, Berliner JL, Ma CB, Feeley BT. Preoperative patient-reported scores can predict postoperative outcomes after shoulder arthroplasty. J Shoulder Elb Surg. 2016;25(6):913–9. https://doi.org/10.1016/j.jse.2016.01.029.
44. 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;45(8):620–5. https://doi.org/10.2519/jospt.2015.5760.
45. Iossifidis A, Ibrahim EF, Petrou C, Galanos A. The development and validation of a questionnaire for rotator cuff disorders: the functional shoulder score. Should Elb. 2015;7(4):256–67. https://doi.org/10.1177/1758573215578589.
46. Negahban H, Behtash Z, Sohani SM, Salehi R. Responsiveness of two Persian-versions of shoulder outcome measures following physiotherapy intervention in patients with shoulder disorders. Disabil Rehabil. 2015;37(24):2300–4. https://doi.org/10.3109/09638288.2015.1005760.
47. Castricini R, Longo UG, De Benedetto M, Loppini M, Zini R, Maffulli N, et al. Arthroscopic-assisted latissimus dorsi transfer for the management of irreparable rotator cuff tears: short-term results. J Bone Joint Surg Am. 2014;96(14):e119—6. https://doi.org/10.2106/JBJS.L.01091.
48. Holmgren T, Oberg B, Adolfsson L, Bjornsson Hallgren H, Johansson K. Minimal important changes in the Constant-Murley score in patients with subacromial pain. J Shoulder Elb Surg. 2014;23(8):1083–90. https://doi.org/10.1016/j.jse.2014.01.014.
49. Koehorst ML, van Trijffel E, Lindeboom R. Evaluative measurement properties of the patient-specific functional scale for primary

shoulder complaints in physical therapy practice. J Orthop Sports Phys Ther. 2014;44(8):595–603. https://doi.org/10.2519/jospt.2014.5133.
50. Lundquist CB, Dossing K, Christiansen DH. Responsiveness of a Danish version of the disabilities of the arm, shoulder and hand (DASH) questionnaire. Dan Med J. 2014;61(4):A4813.
51. Skare O, Mowinckel P, Schroder CP, Liavaag S, Reikeras O, Brox JI. Responsiveness of outcome measures in patients with superior labral anterior and posterior lesions. Should Elb. 2014;6(4):262–72. https://doi.org/10.1177/1758573214534650.
52. van de Water AT, Shields N, Davidson M, Evans M, Taylor NF. Reliability and validity of shoulder function outcome measures in people with a proximal humeral fracture. Disabil Rehabil. 2014;36(13):1072–9. https://doi.org/10.3109/09638288.2013.829529.
53. Kukkonen J, Kauko T, Vahlberg T, Joukainen A, Aarimaa V. Investigating minimal clinically important difference for constant score in patients undergoing rotator cuff surgery. J Shoulder Elb Surg. 2013;22(12):1650–5. https://doi.org/10.1016/j.jse.2013.05.002.
54. 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;8:40. https://doi.org/10.1186/1749-799X-8-40.
55. Christie A, Dagfinrud H, Garratt AM, Ringen Osnes H, Hagen KB. Identification of shoulder-specific patient acceptable symptom state in patients with rheumatic diseases undergoing shoulder surgery. J Hand Ther. 2011;24(1):53–60; quiz 1. https://doi.org/10.1016/j.jht.2010.10.006.
56. 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;20(1):115–28. https://doi.org/10.1123/jsr.20.1.115.
57. Ekeberg OM, Bautz-Holter E, Keller A, Tveita EK, Juel NG, Brox JI. A questionnaire found disease-specific WORC index is not more responsive than SPADI and OSS in rotator cuff disease. J Clin Epidemiol. 2010;63(5):575–84. https://doi.org/10.1016/j.jclinepi.2009.07.012.
58. 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;40(7):413–21. https://doi.org/10.2519/jospt.2010.3209.
59. 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;92(2):296–303. https://doi.org/10.2106/JBJS.H.01296.
60. Mintken PE, Glynn P, Cleland JA. Psychometric properties of the shortened disabilities of the arm, shoulder, and hand questionnaire (QuickDASH) and numeric pain rating scale in patients with shoulder pain. J Shoulder Elb Surg. 2009;18(6):920–6. https://doi.org/10.1016/j.jse.2008.12.015.
61. 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 Elb Surg. 2009;18(6):927–32. https://doi.org/10.1016/j.jse.2009.03.021.
62. Moser JS, Barker KL, Doll HA, Carr AJ. Comparison of two patient-based outcome measures for shoulder instability after non-operative treatment. J Shoulder Elb Surg. 2008;17(6):886–92. https://doi.org/10.1016/j.jse.2008.05.040.
63. Leggin BG, Michener LA, Shaffer MA, Brenneman SK, Iannotti JP, Williams GR Jr. The Penn shoulder score: reliability and validity. J Orthop Sports Phys Ther. 2006;36(3):138–51. https://doi.org/10.2519/jospt.2006.36.3.138.
64. Paul A, Lewis M, Shadforth MF, Croft PR, Van Der Windt DA, Hay EM. A comparison of four shoulder-specific questionnaires in primary care. Ann Rheum Dis. 2004;63(10):1293–9. https://doi.org/10.1136/ard.2003.012088.
65. 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 Elb Surg. 2002;11(6):587–94. https://doi.org/10.1067/mse.2002.127096.
66. Simovitch R, Flurin PH, Wright T, Zuckerman JD, Roche CP. Quantifying success after total shoulder arthroplasty: the substantial clinical benefit. J Shoulder Elb Surg. 2018;27(5):903–11. https://doi.org/10.1016/j.jse.2017.12.014.
67. Michener LA, Snyder Valier AR, McClure PW. Defining substantial clinical benefit for patient-rated outcome tools for shoulder impingement syndrome. Arch Phys Med Rehabil. 2013;94(4):725–30. https://doi.org/10.1016/j.apmr.2012.11.011.
68. Tran G, Dube B, Kingsbury SR, Tennant A, Conaghan PG, Hensor EMA. Investigating the patient acceptable symptom state cut-offs: longitudinal data from a community cohort using the shoulder pain and disability index. Rheumatol Int. 2020;40(4):599–605. https://doi.org/10.1007/s00296-019-04486-3.
69. Chamberlain AM, Hung M, Chen W, Keener JD, McAllister J, Ebersole G, et al. Determining the patient acceptable symptomatic state for the ASES, SST, and VAS pain after total shoulder arthroplasty. J Shoulder Elb Arthroplast. 2017;1:247154921772004. https://doi.org/10.1177/2471549217720042.
70. Sciascia AD, Morris BJ, Jacobs CA, Edwards TB. Responsiveness and internal validity of common patient-reported outcome measures following total shoulder arthroplasty. Orthopedics. 2017;40(3):e513–e9. https://doi.org/10.3928/01477447-20170327-02.
71. O'Halloran B, Wright A, Cook CE. Criterion validation of the rate of recovery, a single alphanumeric measure, in patients with shoulder pain. Int J Sports Phys Ther. 2013;8(6):784–92.
72. 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. https://doi.org/10.1097/BRS.0b013e31815e3a10.
73. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol. 2003;56(5):395–407. https://doi.org/10.1016/s0895-4356(03)00044-1.
74. Beaton DE, van Eerd D, Smith P, van der Velde G, Cullen K, Kennedy CA, et al. Minimal change is sensitive, less specific to recovery: a diagnostic testing approach to interpretability. J Clin Epidemiol. 2011;64(5):487–96. https://doi.org/10.1016/j.jclinepi.2010.07.012.
75. Copay AG, Chung AS, Eyberg B, Olmscheid N, Chutkan N, Spangehl MJ. Minimum clinically important difference: current trends in the orthopaedic literature, part I: upper extremity: a systematic review. JBJS Rev. 2018;6(9):e1. https://doi.org/10.2106/JBJS.RVW.17.00159.
76. Ward MM, Guthrie LC, Alba M. Domain-specific transition questions demonstrated higher validity than global transition questions as anchors for clinically important improvement. J Clin Epidemiol. 2015;68(6):655–61. https://doi.org/10.1016/j.jclinepi.2015.01.028.
77. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. Ann Rheum Dis. 2005;64(1):34–7. https://doi.org/10.1136/ard.2004.023028.
78. Hawkins RJ, Thigpen CA. Selection, implementation, and interpretation of patient-centered shoulder and elbow outcomes. J Shoulder Elb Surg. 2018;27(2):357–62. https://doi.org/10.1016/j.jse.2017.09.022.
79. Roy JS, MacDermid JC, Woodhouse LJ. A systematic review of the psychometric properties of the Constant-Murley score. J Shoulder Elb Surg. 2010;19(1):157–64. https://doi.org/10.1016/j.jse.2009.04.008.
80. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences