



Quantifying the minimal and substantial clinical benefit of the Constant-Murley score and the Disabilities of the Arm, Shoulder and Hand score in patients with calcific tendinitis of the rotator cuff

Jan K.G. Louwerens, MD ^{a,*}, Michel P.J. van den Bekerom, MD, PhD ^b,
Barend J. van Royen, MD, PhD ^c, Denise Eygendaal, MD, PhD ^{c,d},
Arthur van Noort, MD, PhD ^a, Inger N. Sierevelt, MSc ^a

^a Department of Orthopaedic Surgery, Spaarne Gasthuis, Hoofddorp, The Netherlands

^b Department of Orthopaedic Surgery, Onze Lieve Vrouwe Gasthuis (OLVG), Amsterdam, The Netherlands

^c Department of Orthopaedic Surgery, Amsterdam University Medical Centers, Amsterdam, The Netherlands

^d Department of Orthopaedic Surgery, Amphia, Breda, The Netherlands

ARTICLE INFO

Keywords:

Rotator cuff calcific tendinitis
minimal clinically important difference
substantial clinical benefit
minimally invasive treatment
prognostic factors
outcome measures

Level of evidence: Basic Science Study;
Validation of Outcome Instruments

Background: To aid the interpretation of clinical outcome scores, it is important to determine the measurement properties. The aim of this study was to establish the minimal clinically important difference (MCID) and substantial clinical benefit (SCB) for the Constant-Murley score and Disabilities of the Arm, Shoulder and Hand score in patients with long-lasting rotator cuff calcific tendinitis treated with high-energy extracorporeal shockwave therapy and ultrasound guided needling. The secondary purpose was to assess the responsiveness of both questionnaires and to identify variables associated with achieving the MCID and SCB.

Methods: A prospective cohort of 80 patients with rotator cuff calcific tendinitis was analyzed. Two anchor-based methods were used to calculate the MCID and SCB. Effect sizes and standardized response means were calculated to assess the responsiveness. Additional univariate logistic regression analyses were performed to identify factors associated with the achievement of the MCID and SCB.

Results: For the Constant-Murley score, we found an MCID and SCB of 9.8 and 19.9, respectively, based on the mean change method and 5.5 and 10.5, respectively, based on receiver operating characteristic analysis. For the Disabilities of the Arm, Shoulder and Hand score, we found an MCID and SCB of –8.2 and –19.6, respectively, with the former and –11.7 and –12.5, respectively, with the latter. The responsiveness of both outcome measures was good, with large effect sizes and standardized response means. The radiographic resorption after 6 weeks and after 6 months appeared to be the most important positive predictor for achieving the MCID and SCB after 6 months.

Conclusion: This study established the MCID, SCB, and responsiveness for patients with long-lasting rotator cuff calcific tendinitis who were treated with minimally invasive treatment options. With this information, physicians can distinguish between a statistically significant difference and a clinically relevant benefit. Successful radiographic resorption after 6 weeks and after 6 months was associated with achieving clinically significant improvement after treatment.

© 2020 The Authors. Published by Elsevier Inc. on behalf of American Shoulder and Elbow Surgeons. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Calcific tendinitis of the rotator cuff is a painful condition characterized by the deposition of hydroxyapatite crystals in the rotator cuff.³⁷ It is a frequent cause of subacromial pain, and patients experience overhead activity-related pain.²⁴ Although the

condition might be self-limiting in the majority of cases, the symptoms can be severe and prolonged. In the past few years, numerous clinical trials and reviews have been published on the minimally invasive treatment of this condition.^{4,25} In general, these studies have drawn their conclusions based on changes in function, pain, and general health, measured by clinical outcome scores or patient-reported outcome measures. The term “statistical significance” is frequently used to describe a change in outcome of these clinical scores, which does not necessarily mean a relevant benefit for the patient. For example, even small changes can be statistically

The study was approved by the Medical Ethics Review Committee (no. NL44205.094.13) and Institutional Review Board of Spaarne Gasthuis Hospital (no. 2013.26).

* Corresponding author: Jan K.G. Louwerens, MD, Spaarne Gasthuis, Spaarnepoort 1, 2134TM, Hoofddorp, the Netherlands.

E-mail address: jlouwerens@spaarnegasthuis.nl (J.K.G. Louwerens).

<https://doi.org/10.1016/j.jseint.2020.05.001>

2666-6383/© 2020 The Authors. Published by Elsevier Inc. on behalf of American Shoulder and Elbow Surgeons. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

significant in large clinical trials while the real question is whether these changes are clinically relevant for the patient. Therefore, there is an increased need to establish clinical relevance within these outcome measures. The outcome measures should have clearly defined measurement properties such as validity, reliability, and responsiveness. Responsiveness is defined as the ability of an instrument to detect change over time in the construct to be measured.^{27,28} To aid the interpretation of clinical outcome score findings, researchers developed the concept of the minimal clinically important difference (MCID), defined as the smallest change in score in the domain of interest that patients perceive as important.¹⁹ The MCID can help interpret the magnitude of effects of interventions as well as help researchers to determine a more accurate sample size in future studies. An alternative clinically significant measure is the substantial clinical benefit (SCB), defined as the change in outcome associated with patient perception of a large meaningful improvement.²⁹

The distribution- and anchor-based methods are 2 common approaches to calculate the MCID.¹⁰ The distribution-based method uses statistical analysis to determine minimal clinically important changes that occur beyond expected measurement error or variance. The anchor-based approach uses “anchor” questions that aim to evaluate domains such as pain and function to classify changes in clinical outcome scores.

Although systematic reviews addressing the MCID in shoulder outcome scores are available, they report a wide range of MCID values and are investigated in a wide range of shoulder pathology.^{9,14} Even fewer data are available on the SCB values for frequently used shoulder outcome scores.

The main purpose of this study was therefore to establish the MCID and SCB for the Constant-Murley score (CMS) and the Disabilities of the Arm, Shoulder and Hand (DASH) score in patients with long-lasting calcific tendinitis of the rotator cuff treated with high-energy extracorporeal shockwave therapy or ultrasound guided needling. The secondary purpose was to assess the responsiveness of both questionnaires and to identify variables associated with achieving the MCID and SCB.

Materials and methods

Study population

The study population consisted of patients included in a randomized clinical trial evaluating the effect of high-energy shockwave therapy and ultrasound-guided needling for calcific tendinitis of the rotator cuff.²³ The inclusion criteria for this study were as follows: age ≥ 18 years; clinical signs of subacromial pain syndrome for >4 months; standardized radiographs showing a calcific deposit with a diameter ≥ 5 mm in the rotator cuff; and completion of a nonsurgical treatment program that was unsuccessful, including nonsteroidal anti-inflammatory drugs, physiotherapy (centric and eccentric rotator cuff–strengthening exercises in combination with scapular stabilization), and subacromial infiltration with a corticosteroid.

The study was registered in the Netherlands Trial Register (NL4304/NTR4448). Informed consent forms were signed by all participating patients.

Outcome measures

The CMS and a region-specific DASH score were available for all patients at baseline and after 6 months.^{8,18} Outcomes at baseline and 6 months' follow-up were used for MCID and SCB calculation. For this purpose, an anchor question (7-point global transition rating scale) concerning shoulder complaints was added at 6

months' follow-up. Additional baseline characteristics such as age, sex, workload, dominance, and treatment, as well as radiographic parameters (after 6 weeks and after 6 months), were also assessed.

Constant-Murley score

The CMS is a standardized, simple clinical method of assessing shoulder function and has a maximum score of 100 points, with both subjective (35 points) and objective (65 points) components.⁸ The subjective parameters assess the degree of pain perception (15 points) and the ability to perform the normal tasks of daily living in both activity- and position-related terms (20 points). The objective parameters include testing of active range of motion (40 points) and muscle strength (25 points). The CMS has established measurement properties.^{32,40}

DASH score

The DASH outcome measure is a 30-item, self-reported questionnaire designed to measure physical function and symptoms in patients with various musculoskeletal disorders of the upper limb.¹⁸ It has been validated in the Dutch language.³⁹ The DASH score has acceptable measurement properties.^{33,38} The score ranges from 0 (no disabilities) to 100 (most severe disabilities) and is considered incomplete if more than 3 items (10%) are missing.³³

Anchor question

As an external anchor for this study, a 7-point global rating-of-change (GRC) scale was used. Patients were asked a single question to indicate how their symptoms had changed since baseline^{12,19}: “Since the start of the treatment, in what way would you describe the change in symptoms related to your shoulder condition?” The answer options were (1) much improved, (2) improved, (3) slightly improved, (4) unchanged, (5) slightly worse, (6) worse, and (7) much worse. The “slightly improved” and “improved” categories were used to identify patients who experienced minimally important improvement and substantial improvement, respectively.

Statistical analysis

Statistical analysis was performed by use of SPSS software (version 26.0; IBM, Armonk, NY). Summary statistics were used to describe patients' clinical characteristics and score distributions of the CMS and DASH score. Continuous variables were presented as means with standard deviations. Categorical variables were presented as frequencies with accompanying percentages. To assess the suitability of the anchors, Spearman rank correlation coefficients were calculated between the change-from-baseline scores and the anchor. If the Spearman rank correlation coefficient was >0.50 , the anchor was considered suitable.⁶ Because the number of patients who reported a decline in shoulder function was small ($n = 6$), assessment of responsiveness and estimation of the MCID and SCB were only performed for the clinically improved patients.

Responsiveness

For each transition category of the GRC scale, effect sizes and standardized response means were calculated to assess the responsiveness. The effect size and standardized response mean were calculated by dividing the mean change-from-baseline score at 6 months' follow-up by the standard deviation of the baseline score and by the standard deviation of the change-from-baseline score, respectively.^{20,22} Hypotheses for these variables were

formulated for both outcome measures according to the definitions of Cohen,⁷ with absolute values being at least small (0.2), medium (0.5), and large (0.8) for patients reporting slight improvement, improvement, and much improvement, respectively.

MCID and SCB estimation

To calculate the SCB and MCID of the DASH score and CMS, 2 anchor-based methods were applied, using the GRC scale as anchor. The MCID and SCB were calculated as the mean change score (95% confidence interval [CI]) of both outcomes for those patients who reported being slightly improved and being improved on the GRC scale, respectively.¹⁵ Next, the MCID and SCB were estimated using the receiver operating characteristic cutoff points of the DASH change score and CMS change score. The Youden index was used to assess the optimal cutoff points with the smallest number of misclassified patients for each outcome measure. Additional bootstrapping (statistical resampling) procedures (with 1000 bootstrap samples) were performed to estimate the standard error of the retrieved cutoff values and calculate the 95% CI.¹³

The area under the curve was calculated as a measure of accuracy. It represents the probability that patients with and without minimal or substantial improvement are correctly classified (according to the external criterion). This area ranges from 0.5 (accuracy based only on chance) to 1.0 (perfect accuracy). An area under the curve > 0.7 (with a 95% CI lower bound \geq 0.5) was considered a good discriminator.³⁵ The external criterion for SCB was defined as the merged GRC categories of improved and much improved. For the minimal important change, the category of slightly improved was added.

Factors associated with MCID and SCB

Additional univariate logistic regression analyses were performed to identify factors associated with the achievement of the MCID and SCB of the CMS and DASH score derived from the mean change analysis. Odds ratios and 95% CIs were calculated. $P < .05$ was considered statistically significant.

Results

Baseline characteristics

Between May 2014 and December 2017, a total of 82 patients were randomized and treated with either high-energy extracorporeal shockwave therapy or ultrasound-guided needling. After 6 months, 80 patients (97.5%) were available for follow-up. The mean age was 52 ± 9 years, and 51 of patients (64%) were women. Calcifications were predominantly located in the supraspinatus muscle (Table I).

Clinical outcome measures

The overall mean CMS at baseline was 67.3 ± 12.1 , with an improvement after 6 months to 80.5 ± 17.3 . The DASH score at baseline was 36.6 ± 15.9 , which declined after 6 months to 20.9 ± 18.5 . The subgroup scores for each anchor category can be found in Table II.

Anchor

Among 82 included patients, 80 patients filled out the GRC scale after 6 months. Using the GRC scale, 6 patients (8%) reported a deterioration in function (much, $n = 2$; considerable, $n = 1$; and

Table I
Patient and clinical characteristics ($n = 80$)

	Data
Age, mean (SD), yr	52.1 (9.0)
Sex, n (%)	
Male	29 (36)
Female	51 (64)
Dominant arm treated, n (%)	51 (64)
Workload, n (%)	
Light	39 (59)
Medium	16 (24)
Heavy	11 (17)
Unemployed	14
Gärtner classification, n (%)	
I	34 (42)
II	46 (58)
Treatment, n (%)	
H-ESWT	40 (50)
UGN	40 (50)
Location, n (%)	
Supraspinatus	69 (86)
Infraspinatus	7 (9)
Subscapularis	4 (5)
Magnitude at baseline, mean (SD), mm	15.7 (5.2)
Resorption at 6 weeks, n (%)	52 (65)
Resorption at 6 mo, n (%)	61 (76)

SD, standard deviation; H-ESWT, high-energy extracorporeal shockwave therapy; UGN, ultrasound-guided needling.

slight, $n = 3$), 13 (16%) reported no change in function, and 61 (76%) reported improvement (Table II).

Correlation of change in clinical outcome with anchor

Both outcomes were significantly correlated with the anchor, with values of 0.73 and -0.73 , with the absolute values of both coefficients exceeding the threshold of 0.50, indicating that the GRC scale was suitable as an anchor.

Responsiveness

Effect sizes and standardized response means of subgroups formed by the transition GRC scale are presented in Table II. Both variables increased with increased reported improvement on the GRC scale.

All effect sizes and standardized response means of all transition categories met the aforementioned criteria of Cohen.⁷ The effect sizes and standardized response means of patients who did not experience a clinical improvement did not exceed 0.2 (Table III).

MCID and SCB estimation

The MCID and SCB values of the CMS and DASH score based on both methods are presented in Table III. The mean change method generally revealed higher MCID values than did receiver operating characteristic analysis.

Factors associated with MCID and SCB

Radiographic resorption of the calcific deposit after 6 weeks (CMS and DASH score) and 6 months (DASH score) was significantly associated with the achievement of the MCID. For the SCB, resorption after 6 weeks (CMS) and 6 months (DASH score) was associated with achieving the SCB. Detailed results are reported in Tables IV and V.

Table II
Responsiveness of CMS and DASH score: ES and SRM

	Baseline, mean (SD)	6 mo, mean (SD)	Change, mean (95% CI)	ES	SRM
CMS					
No change (n = 13)	67.1 (16.1)	66.8 (14.4)	−0.3 (−4.7 to 4.0)	0.02	0.04
Slightly improved (n = 19)	68.0 (11.8)	77.8 (11.4)	9.8 (3.7 to 15.9)	0.83	0.80
Improved (n = 15)	66.0 (10.5)	85.9 (10.5)	19.9 (13.6 to 26.3)	1.90	1.75
Much improved (n = 27)	69.0 (11.3)	93.3 (6.4)	24.3 (19.3 to 29.4)	2.15	1.91
DASH score					
No change (n = 13)	34.1 (19.1)	32.6 (19.4)	−1.7 (−8.0 to 4.6)	−0.09	−0.17
Slightly improved (n = 19)	36.1 (15.0)	27.9 (12.5)	−8.2 (−14.7 to −7.5)	−0.55	−0.61
Improved (n = 15)	39.0 (15.5)	19.4 (11.7)	−19.6 (−26.3 to −12.8)	−1.26	−1.61
Much improved (n = 27)	35.3 (14.1)	4.7 (3.9)	−30.7 (−36.2 to −25.1)	−2.18	−2.19

SD, standard deviation; CMS, Constant-Murley score; DASH, Disabilities of the Arm, Shoulder and Hand; ES, effect size; SRM, standardized response mean; CI, confidence interval.

Table III
Results of MCID and SCB calculations based on mean change method and cutoff point of ROC curve

	Mean change	ROC (95% CI)	AUC (95% CI)	TP, %	TN, %
CMS					
MCID	9.8 (3.7 to 15.9)	5.5 (1.1 to 9.9)	0.94 (0.89 to 0.99)	87	89
SCB	19.9 (13.6 to 26.3)	10.5 (7.8 to 13.2)	0.89 (0.81 to 0.97)	86	87
DASH					
MCID	−8.2 (−14.7 to −7.5)	−11.7 (−19.1 to −4.4)	0.88 (0.80 to 0.96)	74	94
SCB	−19.6 (−26.3 to −12.8)	−12.5 (−18.8 to −6.2)	0.89 (0.81 to 0.96)	88	78

MCID, minimal clinically important difference; SCB, substantial clinical benefit; ROC, receiver operating characteristic; AUC, area under curve; CI, confidence interval; TP, true positive; TN, true negative; CMS, Constant-Murley score; DASH, Disabilities of the Arm, Shoulder and Hand.

Discussion

The primary purpose of this study was to assess the MCID and SCB of 2 frequently used shoulder metrics in a population of patients with rotator cuff calcific tendinitis. For the CMS, we found an MCID and SCB of 9.8 and 19.9, respectively, based on the mean change method and 5.5 and 10.5, respectively, based on receiver operating characteristic analysis. For the DASH score, we found an MCID and SCB of −8.2 and −19.6, respectively, with the former and −11.7 and −12.5, respectively, with the latter. The responsiveness of both outcome measures was good, with effect sizes and standardized response means that were larger than required in patients experiencing improvement and were small (<0.2) in the unresponsive group. All area-under-the-curve calculations exceeded 0.70 with a lower bound of the 95% CI that was higher than 0.50, indicating adequate responsiveness (Table III). The radiographic resorption after 6 weeks and after 6 months appeared to be the most important positive predictor for achieving the MCID and SCB after 6 months. Assessment of the MCID and SCB is of value to determine whether a statistically significant clinical outcome is also clinically relevant for a specific patient category. It can also help researchers when calculating the sample size for future clinical studies.

An explanation for why the responsiveness was so high could be that the study population was very homogeneous, with small standard deviations across all outcome measurements. The large treatment effect in both groups could also have contributed.

In the past 5 years, numerous articles have been published on the clinical relevance of shoulder outcome scores.^{11,16,17,21,30} Recent reviews have discussed that the MCID often varies widely.^{9,14,34} The range of reported MCID values was broad: 3 to 36 (median estimate, 8.3) for the CMS and −4.5 to 25.4 (median estimate, 10.2) for the DASH score. The patient categories, treatment techniques, and methodologic protocols differed substantially among the included studies.

It is important to realize that the MCID and SCB do not have fixed values. They are influenced by numerous variables, such as the baseline score, patient category, treatment effect, anchor

question, and definition of minimal clinical difference.^{2,3} It is therefore important to calculate the MCID for different patient categories and for different shoulder metrics. Previous attempts to establish a clear relationship between these variables and the MCID values were not successful.¹⁴ In this study, an anchor-based approach was chosen with 2 different statistical methods to calculate the MCID and SCB. The most accurate way to calculate the MCID remains unclear.^{2,31}

This study found that radiographic signs of resorption after 6 weeks and after 6 months were the only variables that were associated with a higher chance of reaching clinical important outcomes. This finding is of great importance for physicians treating patients with rotator cuff calcific tendinitis and helps deal with the patients' expectation management after treatment.

When one is interpreting the MCID and SCB in clinical research, it is important to realize that an outcome measure also has a smallest detectable change (or measurement error), defined as the smallest change in score that one can detect with an instrument. For the purpose of individual monitoring of patients, the smallest detectable change should be smaller than the MCID to be able to distinguish a minimal clinically relevant difference from the measurement error.³⁶ For the DASH score, the reported smallest detectable change in the literature ranges between 7.9 and 16.3, and for the CMS, measurement errors as high as 17 and 23 have been reported.^{1,5,26,38} Although these calculations were performed for different patient categories after different types of treatment, it remains important to differentiate the smallest detectable change from the MCID. In light of our findings, the MCID values for both the DASH score (8.2) and CMS (9.8) might have been clinically relevant for the patients, but it is possible that they cannot be distinguished from the measurement error. Finally, although the MCID might be exceeded by the smallest detectable change on an individual level, it could still be used in larger clinical trials.³⁶

The results of this study must be interpreted in light of several limitations. First, the use of GRC scales has been questioned because such scales are seldom thoroughly investigated in terms of validity and reliability. Whether patients are able to recall their previous

Table IV
Univariate factors associated with minimal or substantial improvement in CMS, based on mean change in CMS

	MCID		SCB	
	OR	P Value	OR	P Value
Age	0.96 (0.91 to 1.01)	.12	1.00 (0.95 to 1.06)	.95
Sex				
Female	Ref		Ref	
Male	0.65 (0.26 to 1.64)	.36	0.54 (0.19 to 1.49)	.23
Workload				
Light	Ref		Ref	
Middle	1.50 (0.45 to 4.99)	.51	1.68 (0.49 to 5.82)	.41
Heavy	2.40 (0.55 to 10.46)	.24	3.36 (0.84 to 13.48)	.09
Resorption at 6 weeks	3.40 (1.30 to 8.89)	.01*	3.12 (1.02 to 9.50)	.046*
Resorption at 6 mo	2.64 (0.91 to 7.66)	.07	3.23 (0.85 to 12.30)	.09
Gärtner classification				
I	Ref		Ref	
II	0.76 (0.31 to 1.87)	.56	0.64 (0.25 to 1.64)	.35
Treatment				
H-ESWT	Ref		Ref	
UGN	1.00 (0.41 to 2.41)	>.999	1.00 (0.39 to 2.55)	>.999
Size of calcific deposit at baseline	0.99 (0.91 to 1.08)	.79	0.96 (0.87 to 1.06)	.34
Dominant arm treated	0.79 (0.32 to 2.00)	.62	1.43 (0.53 to 3.88)	.48

CMS, Constant-Murley score; MCID, minimal clinically important difference; SCB, substantial clinical benefit; OR, odds ratio; Ref, reference value; H-ESWT, high-energy extracorporeal shockwave therapy; UGN, ultrasound-guided needling.

* Statistically significant ($P < .05$).

Table V
Univariate factors associated with minimal or substantial improvement in DASH score, based on mean change in DASH score

	MCID		SCB	
	OR	P Value	OR	P Value
Age	0.99 (0.94 to 1.04)	.66	0.96 (0.91 to 1.01)	.11
Sex				
Female	Ref		Ref	
Male	0.44 (0.17 to 1.18)	.10	0.70 (0.27 to 1.84)	.47
Workload				
Light	Ref		Ref	
Middle	4.87 (0.97 to 24.44)	.05	1.20 (0.36 to 4.03)	.41
Heavy	6.26 (0.72 to 54.41)	.10	8.00 (1.48 to 43.2)	.02*
Resorption 6 weeks	3.39 (1.26 to 9.11)	.02*	2.58 (0.94 to 7.11)	.07
Resorption 6 mo	5.42 (1.83 to 16.05)	.002*	4.64 (1.23 to 17.51)	.02*
Gärtner classification				
I	Ref		Ref	
II	0.68 (0.26 to 1.79)	.43	1.18 (0.47 to 1.96)	.73
Treatment				
H-ESWT	Ref		Ref	
UGN	1.32 (0.51 to 3.41)	.57	0.46 (0.18 to 1.15)	.10
Size of calcific deposits at baseline	0.98 (0.88 to 1.05)	.34	1.01 (0.93 to 1.11)	.75
Dominant arm treated	0.91 (0.34 to 2.43)	.85	1.06 (0.42 to 2.70)	.91

DASH, Disabilities of the Arm, Shoulder and Hand; MCID, minimal clinically important difference; SCB, substantial clinical benefit; OR, odds ratio; Ref, reference value; H-ESWT, high-energy extracorporeal shockwave therapy; UGN, ultrasound-guided needling.

* Statistically significant ($P < .05$).

status has been debated. The scales have shown to be influenced by recent events and the patient's status, as well as the change over time. Owing to these factors, the GRC scale might be correlated more to the post-treatment score than the change-from-baseline score. However, this was not the case in our study. Second, there is no established external criterion for determining the MCID or SCB. In this study, a 7-point global assessment scale was used, but other authors have used a 9- or 11-point GRC scale or a different scale, such as a visual analog scale. Finally, the smallest detectable change was not determined in this study.

Conclusions

This study established the MCID, SCB, and responsiveness for patients with long-lasting rotator cuff calcific tendinitis who were

treated with minimally invasive treatment options. With this information, physicians can distinguish between a statistically significant difference and a clinically relevant benefit. Successful radiographic resorption after 6 weeks and after 6 months was associated with achieving clinically significant improvement after treatment.

Disclaimer

The authors, their immediate families, and any research foundations with which they are affiliated have not received any financial payments or other benefits from any commercial entity related to the subject of this article.

References

- Alotaibi NM, Aljadi SH, Alrowayeh HN. Reliability, validity and responsiveness of the Arabic version of the Disability of Arm, Shoulder and Hand (DASH-Arabic). *Disabil Rehabil* 2016;38:2469–78. <https://doi.org/10.3109/09638288.2015.1136846>.
- Angst F, Aeschlimann A, Angst J. The minimal clinically important difference raised the significance of outcome effects above the statistical level, with methodological implications for future studies. *J Clin Epidemiol* 2017;82:128–36. <https://doi.org/10.1016/j.jclinepi.2016.11.016>.
- Angst F, Benz T, Lehmann S, Aeschlimann A, Angst J. Multidimensional minimal clinically important differences in knee osteoarthritis after comprehensive rehabilitation: a prospective evaluation from the Bad Zurzach Osteoarthritis Study. *RMD Open* 2018;8:e000685. <https://doi.org/10.1136/rmdopen-2018-000685>.
- Arirachakaran A, Boonard M, Yamaphai S, Prommahachai A, Kesprayura S, Kongtharvonskul J. Extracorporeal shock wave therapy, ultrasound-guided percutaneous lavage, corticosteroid injection and combined treatment for the treatment of rotator cuff calcific tendinopathy: a network meta-analysis of RCTs. *Eur J Orthop Surg Traumatol* 2017;27:381–90. <https://doi.org/10.1007/s00590-016-1839-y>.
- Ban I, Troelsen A, Kristensen MT. High inter-rater reliability, agreement, and convergent validity of Constant score in patients with clavicle fractures. *J Shoulder Elbow Surg* 2016;25:1577–82. <https://doi.org/10.1016/j.jse.2016.02.022>.
- Cella D, Hahn EA, Dineen K. Meaningful change in cancer-specific quality of life scores: differences between improvement and worsening. *Qual Life Res* 2002;11:207–21. <https://doi.org/10.1023/A:1015276414526>.
- Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
- Constant CR, Murley AH. A clinical method of functional assessment of the shoulder. *Clin Orthop Relat Res* 1987;214:160–4.
- Copay AG, Eyberg B, Chung AS, Zurcher KS, Chutkan N, Spangehl MJ. Minimum clinically important difference: current trends in the orthopaedic literature, part II: lower extremity: a systematic review. *JBJS Rev* 2018;6:e2. <https://doi.org/10.2106/JBJS.RVW.17.00160>.
- Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 2003;56:395–407. [https://doi.org/10.1016/S0895-4356\(03\)00044-1](https://doi.org/10.1016/S0895-4356(03)00044-1).
- 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 Elbow Surg* 2019;28:939–48. <https://doi.org/10.1016/j.jse.2018.10.013>.
- de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in medicine: a practical guide*. Cambridge, England: Cambridge University Press; 2011.
- Efron B, Tibshirani RJ. *An introduction to the bootstrap*. Boca Raton, FL: Chapman and Hall/CRC; 1994.
- 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:e028777. <https://doi.org/10.1136/bmjopen-2018-028777>.
- Hays RD, Farivar SS, Liu H. Approaches and recommendations for estimating minimally important differences for health-related quality of life measures. *COPD* 2005;2:63–7. <https://doi.org/10.1081/COPD-200050663>.
- Henseler JF, Kolk A, van der Zwaal P, Nagels J, Vliet Vlieland TPM, Nelissen RGH. The minimal detectable change of the Constant score in impingement, full-thickness tears, and massive rotator cuff tears. *J Shoulder Elbow Surg* 2015;24:376–81. <https://doi.org/10.1016/j.jse.2014.07.003>.
- 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;23:1083–90. <https://doi.org/10.1016/j.jse.2014.01.014>.
- Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (Disabilities of the Arm, Shoulder, and Hand). *Am J Ind Med* 1996;29:602–8.
- Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertain the minimal clinically important difference. *Control Clin Trials* 1989;10:407–15.
- Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989;27:S178–89.
- Kukkonen J, Kauko T, Vahlberg T, Joukainen A, Aärimaa V. Investigating minimal clinically important difference for Constant score in patients undergoing rotator cuff surgery. *J Shoulder Elbow Surg* 2013;22:1650–5. <https://doi.org/10.1016/j.jse.2013.05.002>.
- Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care* 1990;28:632–42.
- Louwerens JKG, Sierevelt IN, Kramer ET, Boonstra R, van den Bekerom MPJ, van Royen BJ, et al. Comparing ultrasound-guided needling combined with a subacromial corticosteroid injection versus high-energy extracorporeal shock-wave therapy for calcific tendinitis of the rotator cuff. A randomized controlled trial. *Arthroscopy* 2020;36:1823–33.e1. <https://doi.org/10.1016/j.arthro.2020.02.027>.
- Louwerens JKG, Sierevelt IN, van Hove RP, van den Bekerom MPJ, van Noort A. Prevalence of calcific deposits within the rotator cuff tendons in adults with and without subacromial pain syndrome: clinical and radiologic analysis of 1219 patients. *J Shoulder Elbow Surg* 2015;24:1588–93. <https://doi.org/10.1016/j.jse.2015.02.024>.
- Louwerens JKG, Sierevelt IN, van Noort A, van den Bekerom MPJ. Evidence for minimally invasive therapies in the management of chronic calcific tendinopathy of the rotator cuff: a systematic review and meta-analysis. *J Shoulder Elbow Surg* 2014;23:1240–9. <https://doi.org/10.1016/j.jse.2014.02.002>.
- Mahabier KC, den Hartog D, Theyskens N, Verhofstad MHJ, van Lieshout EMM, Bos PK, et al. Reliability, validity, responsiveness, and minimal important change of the Disabilities of the Arm, Shoulder and Hand and Constant-Murley scores in patients with a humeral shaft fracture. *J Shoulder Elbow Surg* 2017;26:e1–12. <https://doi.org/10.1016/j.jse.2016.07.072>.
- Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539–49. <https://doi.org/10.1007/s11136-010-9606-8>.
- Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63:737–45. <https://doi.org/10.1016/j.jclinepi.2010.02.006>.
- Nwachukwu BU, Chang B, Fields K, Rebolledo BJ, Nawabi DH, Kelly BT, et al. Defining the “substantial clinical benefit” after arthroscopic treatment of femoroacetabular impingement. *Am J Sports Med* 2020;48:901–7. <https://doi.org/10.1177/0363546520902736>.
- 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 Elbow Surg* 2019;28:639–47. <https://doi.org/10.1016/j.jse.2018.09.025>.
- Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61:102–9. <https://doi.org/10.1016/j.jclinepi.2007.03.012>.
- Roy JS, MacDermid JC, Woodhouse LJ. A systematic review of the psychometric properties of the Constant-Murley score. *J Shoulder Elbow Surg* 2010;19:157–64. <https://doi.org/10.1016/j.jse.2009.04.008>.
- Solway S, Beaton DE, McConnell S, Bombardier C. The DASH outcome measure user's manual: Disabilities of the Arm, Shoulder and Hand. *J Hand Ther* 2017;30:30–40.e2. <https://doi.org/10.1016/j.jht.2016.06.008>.
- St-Pierre C, Desmeules F, Dionne CE, Frémont P, MacDermid JC, Roy J-S. Psychometric properties of self-reported questionnaires for the evaluation of symptoms and functional limitations in individuals with rotator cuff disorders: a systematic review. *Disabil Rehabil* 2016;38:103–22. <https://doi.org/10.3109/09638288.2015.1027004>.
- Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42. <https://doi.org/10.1016/j.jclinepi.2006.03.012>.
- Terwee CB, Roorda LD, Knol DL, de Boer MR, de Vet HCW. Linking measurement error to minimal important change of patient-reported outcomes. *J Clin Epidemiol* 2009;62:1062–7. <https://doi.org/10.1016/j.jclinepi.2008.10.011>.
- Unthoff HK, Loehr JW. Calcific tendinopathy of the rotator cuff: pathogenesis, diagnosis, and management. *J Am Acad Orthop Surg* 1997;5:183–91.
- van Kampen DA, Willems W, van Beers LWAH, Castelein RM, Scholtes VAB, 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;14:40. <https://doi.org/10.1186/1749-799X-8-40>.
- Veehof MM, Slegers EJA, van Veldhoven NHM, Schuurman AH, van Meeteren NLU. Psychometric qualities of the Dutch language version of the Disabilities of the Arm, Shoulder, and Hand questionnaire (DASH-DLV). *J Hand Ther* 2002;15:347–54. [https://doi.org/10.1016/s0894-1130\(02\)80006-0](https://doi.org/10.1016/s0894-1130(02)80006-0).
- Vrotsou K, Ávila M, Machón M, Mateo-Abad M, Pardo Y, Garin O, et al. Constant-Murley Score: systematic review and standardized evaluation in different shoulder pathologies. *Qual Life Res* 2018;27:2217–26. <https://doi.org/10.1007/s11136-018-1875-7>.