Minimal clinically important difference of Liverpool Elbow Score in elbow arthroplasty

Karthik Vishwanathan, MS (Orth), DNB (Orth), MRCS, MSc (Orth), MCh (Orth) [a],
Omid Alizadehkhaiyat, MD, PhD [b], Graham J. Kemp, MA, DM, FRCPath, FHEA, CSci, FSB [c],
Simon P. Frostick, MA, DM, FRCSEng, FRCSEd (ad hom), FFSTEd [d,*]

[a] *Shri Krishna Hospital and Pramukhswami Medical College, Karamsad, India*
[b] *School of Health Sciences, Liverpool Hope University, Liverpool, UK*
[c] *Department of Musculoskeletal Biology, Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool, UK*
[d] *Musculoskeletal Science Research Group, Institute of Translational Medicine, University of Liverpool, Liverpool, UK*

ARTICLE INFO

Background: The minimal clinically important difference (MCID) that allows the interpretation of small but meaningful changes after intervention has not been reported for the Liverpool Elbow Score (LES). This study aimed to determine the MCID for the LES in patients undergoing total elbow replacement.
Methods: This observational study is based on preoperative and 1-year postoperative clinical outcome of total elbow replacement (Discovery Elbow System) in 71 patients using the LES. A 4-point Likert-like transition scale was used to evaluate patient satisfaction after total elbow replacement. A combination of distribution-based methods (standard deviation [SD] of change in the LES, standard error of mean, smallest detectable change [SDC]) and anchor-based methods (receiver operating curve, difference of mean of change in LES) was used to determine range of MCID values.
Results: The mean change in the LES value was 2.4 (SD, 2.1). The estimated SDC value with upper limit of 90% confidence interval was 1.5. The mean change in LES of "satisfied" and "somewhat satisfied" patient groups was 2.4 (SD, 2.1) and 1.1 (SD, 1.4), respectively, and the difference between both means (MCID based on difference of mean in 2 subgroups) was 1.3. According to receiver operating curve analysis, the value of MCID was 1.6.
Conclusion: The MCID value for the LES was estimated to range between 0.7 and 1.8. The estimated SDC value was 1.5. We propose that the "true" MCID value of the LES would be between 1.6 and 1.8 to ensure that the value is higher than the measurement error of the LES.

The Liverpool Elbow Score (LES) is a region-specific outcome score that is completed by both the clinician and the patient. Validation study by the developers of the score demonstrated that the LES is valid, reliable, and responsive to change in clinical condition of the patient in different elbow conditions.[30] The LES has demonstrated satisfactory responsiveness after total elbow replacement arthroplasty.[39] The LES has found support from other research groups and has been used by independent groups other than the developers of the score. The LES has been used in assessment of functional outcome after total elbow arthroplasty (TEA) in arthritis,[1,3] functional outcome after arthroscopic arthrolysis and hyaluronan gel in

post-traumatic elbow stiffness,[25] difference in functional outcome in a randomized controlled trial comparing platelet-rich plasma and autologous whole blood in chronic tennis elbow,[34] functional outcome after internal fixation of severe olecranon fracture,[23] and functional outcome after internal fixation of intra-articular distal humerus fracture.[26]

Establishing the minimal clinically important difference (MCID) for clinical outcomes scores is an important component of outcomes research to understand treatment effectiveness, particularly from the patient's perspective. The MCID is defined as the smallest change in the value of an outcome instrument that patients perceive as important, beneficial, or harmful.[20] In other words, MCID value differentiates patients who improve from those who do not improve after a therapeutic intervention.[18] The concept of MCID assists in differentiating statistical significance from clinical significance. A statistical test might reveal a significant difference between preoperative and postoperative scores of an outcome instrument; however, if the difference is lower than the MCID value of the outcome instrument, this statistically significant difference is not

deemed to be clinically significant. The MCID value is also helpful in evaluating cost-effectiveness, estimating appropriate sample size for randomized controlled trials, and evaluating power of a nonrandomized study.[20]

To our knowledge, the MCID value has not been determined for the LES. Hence, this study aimed to critically evaluate the MCID of the LES in a large cohort of patients who underwent TEA for various underlying pathologic processes.

Methods

A prospective database of patients who had undergone TEA using the Discovery Elbow System (Biomet Inc, Warsaw IN, USA) was reviewed to identify patients with completed preoperative LES and 1-year postoperative LES and satisfaction questionnaire. Identified patients had undergone TEA for degenerative arthritis (osteoarthritis, post-traumatic arthritis), inflammatory arthritis (rheumatoid arthritis, hemophiliac arthropathy, and psoriatic arthritis), comminuted distal humerus fracture, and loosening of previous elbow prostheses using the Discovery prosthesis between April 2003 and March 2013. Patient demographics are presented in the Results section. All identified cases (N = 71) were operated on and followed up in a single upper limb center.

Outcome assessment

Clinical and functional outcome after TEA was assessed using the LES. Before the operation and at 1-year follow-up, patients first completed the patient-answered questionnaire of the LES (PAQ-LES), followed by completion of the clinical assessment part of the score (CAS-LES) by independent research fellows. PAQ-LES includes 9 questions representing domains of pain (1 question), functional ability to do activities of daily living (7 questions), and functional ability to participate in sporting and recreational activities (1 question). These questions are answered on a 5-point adjectival scale from 0 (maximum disability) to 4 (no functional disability). The CAS-LES includes assessment of range of motion (4 items), muscle strength (1 item), and ulnar nerve function (1 item). The points from PAQ-LES and CAS-LES are then entered individually in a mathematical formula to determine the total LES. In this scoring system, 0 and 10 points indicate the worst and best outcome, respectively.[30,31]

As there is no "gold standard" external criterion to assess change and improvements in the clinical condition of the patient, a 4-point Likert-like transition scale was used to evaluate patient satisfaction after TEA. The options on this scale were very satisfied, satisfied, somewhat satisfied, and unsatisfied. Patients answered this question at postoperative follow-ups.

Estimation of MCID

There is no gold standard method to measure MCID, which can be estimated using either anchor-based or distribution-based methods. It has been recommended that studies use both anchor-based and distribution-based methods to give range of values for MCID and finally triangulate to converge on possible MCID value.[7,42] It is also suggested that anchor-based methods be given greater weight than distribution-based methods for converging on a single value or to narrow the range of possible MCID values. Distribution-based methods are solely used only when suitable external anchors have not been used or are not available for use.[28]

This study determined the MCID using both anchor-based and distribution-based methods. Patient satisfaction was used as a global transition external anchor. This is in accordance with a similar approach by previous studies to estimate clinically meaningful change.[6,32] Adjusting for the change in unsatisfied patients, the MCID can be calculated as the mean change score for satisfied patients

minus the mean change score for somewhat satisfied patients.[6,14,32,37] Receiver operating curve (ROC) analysis is then used to evaluate the point that is closest to the upper left-hand corner of the curve representing MCID.[37,40] A diagonal is drawn from the upper left corner of the ROC to the lower right corner. The point at which this diagonal intersects the curve is considered to be the point closest to the upper left corner, and hence the value of change in the LES at this site of intersection represents the MCID.[37] For ROC analysis, patients who were unsatisfied and somewhat satisfied were grouped into the "not improved" group, and those who were satisfied and very satisfied were grouped into the "improved" group. The entire cohort was included in the ROC analysis rather than just the values adjacent to the point of dichotomy, as this has been shown to increase precision of MCID estimation.[36] Sensitivity is the proportion of patients who are definitely satisfied and whose change in LES is above the threshold MCID value. Specificity is the proportion of patients who are not definitely satisfied and whose change in LES is below the threshold MCID value.

For the distribution-based approach, we first estimated the standard error of mean (SEM) and smallest detectable change (SDC). It has been reported that estimates based on measurement precision of outcome measurement (SEM) are better than estimates based on sample variation (effect size) or those based on statistical significance (paired *t*-test).[9] SEM is an indicator of random error during single use of an outcome instrument and is believed generally to be stable across different populations and different studies.[9,10] SDC or minimum detectable change (MDC) refers to the smallest change in the value of an outcome instrument that is greater than random measurement error associated with use of the instrument.[10] Both SEM and SDC are determined in a stable subgroup of patients in the study cohort. These patients either have perceived no change in clinical condition after an intervention or have experienced negligible or minimal change in their clinical condition.

Repeated application of an outcome instrument in the same patient should give a similar value if the condition has remained stable with no change. However, this is infrequently seen, and more often repeated application gives rise to slight changes in the value of the outcome instrument. This minimum change in value is likely to be due to the measurement error of the outcome instrument. SDC or MDC represents the threshold value beyond which any increase in the score of the outcome instrument is likely to indicate "true" change in clinical condition instead of error due to repeated administration of the outcome tool. A change in the value of an outcome instrument lower than the value of SDC might not indicate true change in the clinical condition, as this is likely to be due to the measurement error of the outcome instrument.

In this study, SEM was calculated as SEM = [standard deviation (SD) of baseline preoperative LES] × [square root of 1 – α], wherein α represents the reliability coefficient of the test-retest value of the outcome instrument in a stable group of patients[5,10,17,19]; α can be represented as either Cronbach α or the intraclass correlation coefficient. Commonly, 90% confidence limit is chosen for MDC and is calculated as $MDC_{90} = (1.65) \times (\text{square root of } 2) \times (SEM)$.[5,10,17,19] Cronbach α based on standardized items was used to measure the reliability coefficient of test-retest in a stable group of patients.[19] Based on patients' response to the 4-point Likert-like satisfaction scale, those patients who felt somewhat satisfied after the TEA were considered to be stable patients, as they probably did not have significant change in their clinical condition. Various threshold values have been reported for the estimation of clinically meaningful change based on SEM including 1 SEM,[41] 1.96 SEM, and 2.77 SEM.[22] Norman et al[24] observed that a value of half the SD of the change in score of the outcome instrument was equal to MCID in a variety of studies, although it is believed that this is a conservative estimate of MCID. SPSS version 18 (SPSS Inc., Chicago, IL, USA) was used to do the statistical analysis.

Results

Study participants

The study included 71 patients who had the required preoperative and postoperative data on LES and patient satisfaction (there were no missing data). The mean age of patients was 64 years (range, 22-93 years); 49 cases were female (69%), and 22 cases were male (31%). There were 52 (73%) and 19 (27%) primary and revision TEA cases, respectively; 45 cases (63%) presented with unilateral elbow and 26 cases (37%) with bilateral elbow involvement. Underlying disease was rheumatoid arthritis in 22 cases (31%), loosening of elbow prosthesis in 19 cases (27%), post-traumatic arthritis in 11 cases (16%), primary osteoarthritis in 11 cases (16%), fracture of distal humerus in 5 cases (7%), hemophiliac arthropathy in 2 cases (3%), and arthritis due to synovial chondromatosis in 1 case (1%).

LES and patient satisfaction

The mean preoperative LES improved from 3.7 (SD, 1.8; range, 0.2-8.3) to 6.1 (SD, 1.71; range, 2.5-9.2). After TEA, 40 patients (56%) were very satisfied, 18 patients (25%) satisfied, 10 patients (14%) somewhat satisfied, and 3 patients (4%) unsatisfied.

Distribution-based approach

MCID

Based on SD of change in LES, the mean change in the value of the LES was 2.4 (range, –1.7 to 8.7). The value of SD for change in LES was 2.13. Using the criterion of $\frac{1}{2} \times$ SD, the estimated MCID was 1.1.

SDC and SEM

Ten patients chose the somewhat satisfied response at 1-year follow-up. These patients were considered to have a stable clinical condition as they did not have significant satisfaction. In these patients, the mean preoperative LES was 3.8 (SD, 2.1; range, 0.2-6.8), and the mean postoperative LES was 4.9 (SD, 1.6; range, 2.7-7.0). Cronbach α based on standardized items was estimated as 0.87. The SD of the baseline preoperative LES of the entire cohort of 71 patients was 1.8. The SEM was estimated to be 0.66, and using criteria of 1 SEM, 1.96 SEM, and 2.77 SEM resulted in the values of 0.7, 1.3, and 1.8, respectively. Based on the mathematical formula already described in the Methods section, the SDC with upper limit of 90% confidence interval was estimated as 1.5.

Anchor-based approach

MCID based on the difference of mean of change in LES

The mean change in LES of the satisfied group was 2.4 (SD, 2.1; range, –1.5 to 6.2), whereas the mean change in LES of the somewhat satisfied group was 1.1 (SD, 1.4; range, –0.7 to 3.5). The difference between both the means (MCID based on difference of mean in 2 subgroups) was 1.3.

MCID based on ROC analysis

ROC analysis revealed the value of change of LES closest to the upper left-hand corner of the graph (MCID) as 1.6 (sensitivity value of 0.69 and specificity value of 0.69) (Fig. 1). The area under the curve was 0.74 (*P* = .007; 95% confidence interval, 0.61-0.88).

Triangulation approach: combined distribution- and anchor-based approaches

Thus, using various methods, it was estimated that the true value of MCID for the LES ranged from minimum value of 0.7 to maximum
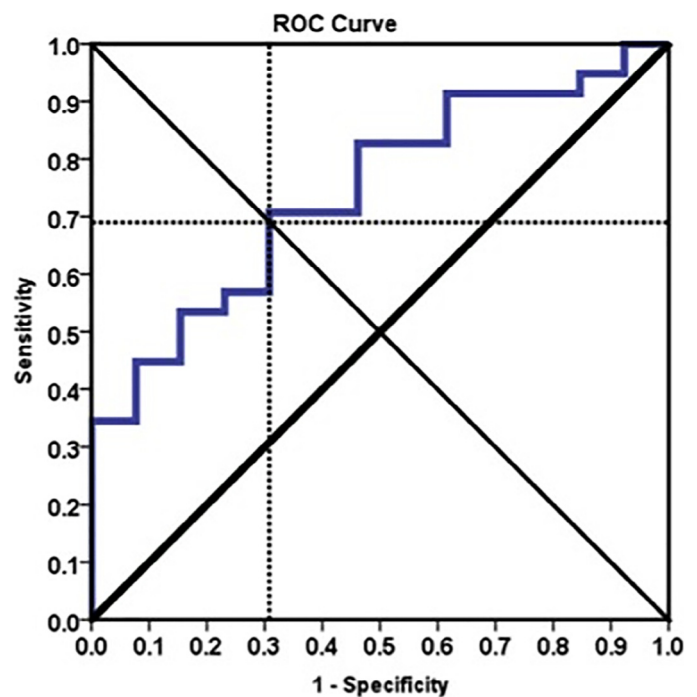

Figure 1 Estimation of minimal clinically important difference using receiver operating curve (*ROC*) analysis. A *diagonal* is drawn from the upper left corner to the lower right corner. The point at which this diagonal intersects the curve represents the point closest to the upper left corner, and hence this represents the minimal clinically important difference value of 1.6. The *vertical and horizontal reference lines* are drawn to x-axis and y-axis, respectively. The *vertical dotted reference line to x-axis* is through a point traversing 1 – specificity of 0.308. The *horizontal dotted reference line to y-axis* is through a point traversing sensitivity of 0.690.

value of 1.8. The SDC value was estimated to be 1.5 in this study. It is proposed that the true MCID value of the LES would be between 1.6 and 1.8 to ensure that the value is higher than the measurement error of the LES.

Discussion

This study has estimated the range of possible values of MCID using a combination of both anchor-based and distribution-based methods. Some of the values were lower than the SDC values, whereas some were higher than the SDC value of 1.5. Most studies have used ROC analysis to estimate MCID, and in our study, the MCID using ROC analysis was 1.6. We propose using the value of 1.6 for calculating sample size for research trials as it is the value just higher than the SDC value.

Patient satisfaction has been previously used as a global transition external criterion to estimate clinically meaningful change.[6,32] It is recommended that correlation of change in value of the outcome score with the external anchor must be at least 0.30 for it to be useful for evaluation of MCID.[27] A previously published study[39] has shown positive correlation between change in value of LES and patient satisfaction (correlation coefficient of 0.35). This justifies our choice of using patient satisfaction as an external anchor.

Some authors have used the 90% upper confidence limit to estimate the SDC, calculated as the product of SEM with square root value of 2 and 1.645.[5,17] These authors have recommended using 90% upper confidence limit of the SDC instead of the 95% upper confidence limit because of its higher precision.[5] By contrast, some authors have used the 95% upper confidence limit to estimate SDC, calculated as the product of SEM with square root values of 2 and 1.96.[16,38] SEM in the stable group of patients has been determined by the

formula SD divided by square root of 2, wherein SD was the change in value of the outcome instrument in a stable group of patients. MDC was determined by obtaining the product of SEM with square root values of 2 and 1.65. MDC calculated by this method represented the upper 90% confidence interval of the MDC. Calculating MDC using this technique tended to overestimate the value of SDC.[8] There is still no consensus regarding whether to use 90% or 95% upper confidence limit for calculating SDC.

The results of this study concur with the findings of de Boer et al,[13] who used patient satisfaction after elbow surgery as an external anchor. In that study, a larger difference in the LES value was associated with higher satisfaction; a smaller difference was observed in patients with lower satisfaction. This study used a 4-point Likert-like scale, whereas de Boer et al[13] used a 10-cm visual analog scale (VAS) to assess patient satisfaction, converted into ordinal data by arbitrarily categorizing patients scoring 0-2.5 points as dissatisfied, 2.5-5 points as somewhat satisfied, 5-7.5 points as moderately satisfied, and 7.5-10 points as very satisfied. This classification could be criticized as patients rating 2.5 points on the VAS could theoretically end in either the dissatisfied or somewhat satisfied group, and similarly, patients rating 5 points on the VAS could end in either the somewhat satisfied or moderately satisfied group. Moreover, the authors did not report the results of satisfaction level, including the proportion of patients in various groups of satisfaction level. It would also have facilitated the data interpretation if the results had been published in the form of mean, SD, and range for various satisfaction level groups.

de Boer et al[13] used patient satisfaction as one of the external anchors and calculated the MCID by estimating the mean change in value of patients classified to be somewhat satisfied. A second external anchor used in that study was the patient's global perceived effect of intervention on a 5-point Likert scale (much improved, slightly improved, no change, slightly worsened, much worsened). For ROC analysis, the much improved response was categorized as the improved group; those who experienced slight improvement, no change, and slight worsening were categorized as the "no change" group. One case of 25 patients who selected the much worsened option was excluded from the analysis. It is recommended that for ROC analysis, all groups should be included and no group should be excluded from the analysis.[36] The study by de Boer et al[13] relied solely on the anchor-based method for MCID calculation, whereas SDC and SEM values were not calculated. It is recommended that authors should use both anchor-based and distribution-based methods to estimate MCID.[36] Moreover, assessment of MCID was not at a uniform interval after the intervention; although the mean follow-up period was 7 months, the assessment was done at varying time intervals ranging from 2 months postoperatively to 15 months after the index procedure.

Dawson et al[12] calculated the MCID of the Oxford Elbow Score (OES) in 74 patients who underwent different elbow operations. It is unclear how many of these patients had TEA. Although SDC of the OES pain subscale was smaller than MCID, the SDC of the OES function and social-psychological subscales was higher than the MCID values of corresponding subscales, and hence the results of the OES must be interpreted with caution. The authors had calculated SDC using 90% confidence upper limit of SEM. As the main intended benefits of TEA in elbow conditions are both pain relief and improved function, it is crucial that both pain and functional subscales show lower values of SDC compared with MCID. Postoperative assessment was done at a similar interval of 6 months after the intervention, and the authors used both anchor-based and distribution-based methods to estimate the MCID. Test-retest reliability was estimated in a sample of 52 subjects with a delay of 48 hours between the readings.[11] In this study, test-retest reliability was estimated in a smaller sample of 10 subjects with longer time duration of 12 months between the readings. This could be a possible explanation for the lower value of the reliability coefficient observed in our study. Higher value of the reliability coefficient is likely to lead to lowering of the value of SDC as it depends on the value of the reliability coefficient. It is also possible that a larger sample size would show lower variability in the LES values.

In the study by Dawson et al,[12] the authors did not calculate the value of the reliability coefficient for the Disabilities of the Arm, Shoulder, and Hand and chose to impute the reliability coefficient from an earlier study by Beaton et al.[4] This approach of using previously published values of the reliability coefficient to estimate MCID was also used by Angst et al.[2] The original validation study on the LES showed that the value of the reliability coefficient was 0.93.[30] If we had used this value of reliability coefficient, the value of SDC would be 1.1. However, we chose to calculate the reliability coefficient in this study as the patient population and intervention administered are different from those in the initial validation study.

A recent review[29] on various outcome instruments for TEA comparing Mayo Elbow Performance Score (MEPS), Hospital for Special Surgery scoring system, Hospital for Special Surgery Total Elbow scoring system, Elbow Functional Assessment, American Shoulder and Elbow Surgeons elbow assessment form, and LES found MEPS the most commonly reported elbow score. The MCID of MEPS and American Shoulder and Elbow Surgeons elbow assessment score has not been reported so far. The authors concluded that comparison of psychometric properties of the elbow scores failed to show any major difference between the outcome instruments.

We agree with the view of Riedel and Beaton[29] that a single total value of outcome instrument should be presented instead of several values for various subdomains like pain, function, and psychosocial components. Another review[35] observed that measurement error of the LES has not been reported. This study addresses the issue of measurement error by describing SEM and SDC.

One of the strengths of this study is the demonstration of satisfactory interpretability of the psychometric property of the LES. Interpretability is defined as the ability to assign and to infer qualitative information from quantitative data.[21] There is no consensus regarding the ideal way to check interpretability of an outcome measure. Terwee et al[33] gave a positive rating if the outcome score was presented as mean and SD in at least 4 subgroups of participants and if values of MCID or minimal important change were presented. This study has described MCID values for the LES. Eechaute et al[15] considered that interpretability of an outcome measure could be inferred if any 2 of the following 4 criteria were achieved: scores presented as mean and SD, comparison of data in various subgroups, evaluation of correlation between change in score and patient's global assessment of change, and correlation between change in score and other commonly used outcome measures. Another study[39] described the LES in 4 subgroups of patients based on patient satisfaction after the total elbow replacement. The mean change in LES and the SD of the change score were estimated for unsatisfied, somewhat satisfied, satisfied, and very satisfied patients. The mean change in LES increased as patient satisfaction improved.

This study has its limitations. The value of the reliability coefficient was 0.85, which was lower than the reliability coefficient (0.93) reported for the initial validation study.[30,31] Possible explanations for this could be the difference in the outcome evaluation time. It is likely that evaluation at 1 year after the index procedure was long. Test-retest validity in the original validation study was performed between 1 and 3 days after the initial evaluation. It is recommended that the value of the test-retest reliability coefficient should be >0.70 to be considered satisfactory.[15] Our value is still acceptable as it is higher than this recommended value.

It is unknown whether the MCID is disease specific or intervention specific. It is possible that the MCID of the LES might be different in rheumatoid arthritis, which usually affects multiple joints, and

in post-traumatic arthritis, wherein the elbow arthroplasty would eradicate the condition of the single joint. Future studies could investigate whether MCID is different for type of surgical interventions (primary and revision TEA) or type of pathologic process (inflammatory and noninflammatory elbow arthritis).

Conclusion

The MCID value for the LES is expected to be between 0.7 and 1.8. As the SDC value was estimated as 1.5, we recommend using the MCID value of 1.6, which is greater than the value of SDC and hence more likely to represent a true change in the clinical condition of the patient.

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

1. Amirfeyz R, Blewitt N. Mid-term outcome of GSB-III total elbow arthroplasty in patients with rheumatoid arthritis and patients with post-traumatic arthritis. Arch Orthop Trauma Surg 2009;129:1505-10. http://dx.doi.org/10.1007/s00402-009-0876-y

2. Angst F, Goldhahn J, Drerup S, Kolling C, Aeschlimann A, Simmen BR, et al. Responsiveness of five outcome measurement instruments in total elbow arthroplasty. Arthritis Care Res 2012;64:1749-55. http://dx.doi.org/10.1002/acr.21744

3. Ashmore AM, Gozzard C, Blewitt N. Use of the Liverpool Elbow Score as a postal questionnaire for the assessment of outcome after total elbow arthroplasty. J Shoulder Elbow Surg 2007;16(Suppl):S55-8. http://dx.doi.org/10.1016/j.jse.2006.08.008

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

5. Beninato M, Portney LG. Applying concepts of responsiveness to patient management in neurologic physical therapy. J Neurol Phys Ther 2011;35:75-81. http://dx.doi.org/10.1097/NPT.0b013e318219308c

6. Bessette L, Sangha O, Kuntz KM, Keller RB, Lew RA, Fossel AH, et al. Comparative responsiveness of generic versus disease-specific and weighted versus unweighted health status measures in carpal tunnel syndrome. Med Care 1998;36:491-502.

7. Cella D, Eton DT, Lai JS, Peterman AH, Merkel DE. Combining anchor and distribution based methods to derive minimal clinically important differences on the Functional Assessment of Cancer Therapy anemia and fatigue scales. J Pain Symptom Manage 2002;24:547-61. http://dx.doi.org/10.1016/S0885-3924(02)00529-8

8. Cleland JA, Whitman JM, Houser JL, Wainner RS, Childs JD. Psychometric properties of selected tests in patients with lumbar spinal stenosis. Spine 2012;12:921-31. http://dx.doi.org/10.1016/j.spinee.2012.05.004

9. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol 2003;56:395-407. http://dx.doi.org/10.1016/S0895-4356(03)00044-1

10. Dawson J, Boller I, Doll H, Lavis G, Cooke P, Jenkinson C. Minimally important change was estimated for the Manchester-Oxford Foot Questionnaire after foot/ankle surgery. J Clin Epidemiol 2014;67:697-705. http://dx.doi.org/10.1016/j.jclinepi.2014.01.003

11. Dawson J, Doll H, Boller I, Fitzpatrick R, Little C, Rees J, et al. The development and validation of a patient reported questionnaire to assess outcomes of elbow surgery. J Bone Joint Surg Br 2008;90-B:466-73. http://dx.doi.org/10.1302/0301-620X.90B4.20290

12. Dawson J, Doll H, Boller I, Fitzpatrick R, Little C, Rees J, et al. Comparative responsiveness and minimal change for the Oxford Elbow Score following surgery. Qual Life Res 2008;17:1257-67. http://dx.doi.org/10.1007/s11136-008-9409-3

13. de Boer YA, Hazes JMW, Winia PCA, Brand R, Rozing PM. Comparative responsiveness of four elbow scoring instruments in patients with rheumatoid arthritis. J Rheumatol 2001;28:2616-23.

14. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. Control Clin Trials 1991;12:142-58.

15. Eechaute C, Vaes P, Van Aerschot L, Asman S, Duquet W. The clinimetric qualities of patient-assessed instruments for measuring chronic ankle instability: a systematic review. BMC Musculoskelet Disord 2007;18:1-11. http://dx.doi.org/10.1186/1471-2474-8-6

16. Gärtner FR, Nieuwenhuijsen K, Sluiter JK. Interpretability of change in the Nurses Work Functioning Questionnaire: minimal important change and smallest detectable change. J Clin Epidemiol 2012;65:1337-47. http://dx.doi.org/10.1016/j.jclinepi.2012.06.013

17. Haley SM, Fragala-Pinkham MA. Interpreting change scores of tests and measures used in physical therapy. Phys Ther 2006;86:735-43.

18. Irrgang JJ, Lubowitz JH. Measuring arthroscopic outcome. Arthroscopy 2008;24:718-22. http://dx.doi.org/10.1016/j.arthro.2007.10.007

19. Kim JK, Park ES. Comparative responsiveness and minimal clinically important differences for idiopathic ulnar impaction syndrome. Clin Orthop Relat Res 2013;471:1406-11. http://dx.doi.org/10.1007/s11999-013-2843-8

20. Kirkley A, Griffin S. Development of disease-specific quality of life measurement tools. Arthroscopy 2003;19:1121-8. http://dx.doi.org/10.1016/j.arthro.2003.10.028

21. Lohr KN, Aaronson NK, Alonso J, Burnam MA, Patrick DL, Perrin EB, et al. Evaluating quality of life and health status instruments: development of scientific review criteria. Clin Ther 1996;18:979-92.

22. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res 1995;4:293-307.

23. Munoz-Mahamud E, Fernandez-Valencia JA, Riba J. Plate osteosynthesis for severe olecranon fractures. J Orthop Surg (Hong Kong) 2010;18:80-4.

24. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. Med Care 2003;41:582-92. http://dx.doi.org/10.1097/01.MLR.0000062554.74615.4C

25. Pederzini LA, Milandri L, Tosi M, Prandini M, Nicoletta F. Preliminary clinical experience with hyaluronan anti-adhesion gel in arthroscopic arthrolysis for post traumatic elbow stiffness. J Orthop Traumatol 2013;14:109-14. http://dx.doi.org/10.1007/s10195-013-0229-z

26. Reising K, Hauschild O, Strohm PC, Suedkamp NP. Stabilisation of articular fractures of the distal humerus: early experience with a novel perpendicular plate system. Injury 2009;40:611-7. http://dx.doi.org/10.1016/j.injury.2008.12.018

27. 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. http://dx.doi.org/10.1016/j.jclinepi.2007.03.012

28. Revicki DA, Erickson PA, Sloan JA, Dueck A, Guess H, Santanello NC. Interpreting and reporting results based on patient-reported outcomes. Value Health 2007;10(Suppl 2):S116-24. http://dx.doi.org/10.1111/j.1524-4733.2007.00274.x

29. Riedel K, Beaton DE. Update on the state of outcome measurement in total elbow arthroplasty research: identifying a need for consensus. J Bone Joint Surg Am 2013;95:e97, (1-8). http://dx.doi.org/10.2106/JBJS.K.01420

30. Sathyamoorthy P, Kemp GJ, Rawal A, Rayner V, Frostick SP. Development and validation of an elbow score. Rheumatology 2004;43:1434-40. http://dx.doi.org/10.1093/rheumatology/keh367

31. Sathyamoorthy P, Kemp GJ, Rawal A, Rayner V, Frostick SP. Development and validation of an elbow score. Rheumatology 2005;44:1081.

32. Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition specific and health status measures in degenerative lumbar spinal stenosis. J Clin Epidemiol 1995;48:1369-78.

33. Terwee CB, Bot SDM, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34-42. http://dx.doi.org/10.1016/j.jclinepi.2006.03.012

34. Thanasas C, Papadimitriou G, Charalambidis C, Paraskevopoulos I, Papanikolaou A. Platelet-rich plasma versus autologous whole blood for the treatment of chronic lateral elbow epicondylitis: a randomized controlled clinical trial. Am J Sports Med 2011;39:2130-4. http://dx.doi.org/10.1177/0363546511417113

35. The B, Reininga IH, El Moumni M, Eygendaal D. Elbow-specific clinical rating systems: extent of established validity, reliability, and responsiveness. J Shoulder Elbow Surg 2013;22:1380-94. http://dx.doi.org/10.1016/j.jse.2013.04.013

36. Turner D, Schünemann HJ, Griffith LE, Beaton DE, Griffiths AM, Critch JN, et al. Using the entire cohort in the receiver operating characteristic analysis maximizes precision of the minimal important difference. J Clin Epidemiol 2009;62:374-9. http://dx.doi.org/10.1016/j.jclinepi.2008.07.009

37. Turner D, Schünemann HJ, Griffith LE, Beaton DE, Griffiths AM, Critch JN, et al. The minimal detectable change cannot reliably replace the minimal important difference. J Clin Epidemiol 2010;63:28-36. http://dx.doi.org/10.1016/j.jclinepi.2009.01.024

38. 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. http://dx.doi.org/10.1186/1749-799X-8-40

39. Vishwanathan K, Alizadehkhaiyat O, Kemp GJ, Frostick SP. Responsiveness of the Liverpool Elbow Score in elbow arthroplasty. J Shoulder Elbow Surg 2013;22:312-7. http://dx.doi.org/10.1016/j.jse.2012.09.003

40. Ward MM, Marx AS, Barry NN. Identification of clinically important changes in health status using receiver operating characteristic curves. J Clin Epidemiol 2000;53:279-84.

41. Wolinsky F, Wan G, Tierney W. Changes in the SF-36 in 12 months in a sample of disadvantaged older adults. Med Care 1998;36:1589-98.

42. Yost KJ, Eton DT. Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. Eval Health Prof 2005;28:172-91. http://dx.doi.org/10.1177/0163278705275340