



Published in final edited form as:

J Shoulder Elbow Surg. 2021 September ; 30(9): e602–e609. doi:10.1016/j.jse.2020.11.027.

The Critical Difference in the DASH Outcome Measure After Essential Upper Extremity Tumor Surgery

Koichi Ogura, MD, PhD¹, Mohamed A. Yakoub, MD¹, Alexander B. Christ, MD¹, Tomohiro Fujiwara, MD, PhD¹, Zarko Nikolic, MD¹, Patrick J. Boland, MD¹, Edward A. Athanasian, John H. Healey, MD^{1,*}

¹Department of Surgery, Orthopaedic Service, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Abstract

Background: The DASH (Disabilities of the Arm, Shoulder and Hand) is a scored questionnaire that is widely used to evaluate the health-related quality-of-life of patients with upper-limb musculoskeletal disorders. However, numerical changes in the measure scores lack clinical significance without meaningful threshold change values of outcome measures that are diagnostically specific. The minimal clinically important difference is useful for the interpretation of scores by defining the smallest change that a patient would perceive. However, the minimal clinically important differences of the scores in orthopaedic oncology patients has not been reported. We aimed to determine the minimal clinically important differences of the measure in orthopaedic oncology patients.

Methods: Data from our health-related quality-of-life Database from 1999 to 2005 were retrospectively reviewed after institutional review board approval. Seventy-eight patients who underwent surgery and completed two surveys during postoperative follow-up were evaluated. Two different methods were used to estimate the minimal clinically important differences:

*Corresponding author: John H. Healey, Chief, Department of Surgery, Orthopaedic Service, Memorial Sloan Kettering Cancer Center, Professor of Orthopaedic Surgery, Weill Cornell Medical College, 1275 York Avenue, New York, NY 10065, USA, TEL: +1-212-639-7610; FAX: +1-212-717-3573; healeyj@mskcc.org.

1. Conceptualization: P. Boland, E. Athanasian, J. Healey
2. Data Curation: K. Ogura, M. Yakoub, A. Christ, T. Fujiwara, Z. Nikolic
3. Formal Analysis: K. Ogura, T. Fujiwara, Z. Nikolic
4. Funding Acquisition: J. Healey
5. Investigation: K. Ogura, M. Yakoub, A. Christ, T. Fujiwara, Z. Nikolic
6. Methodology: K. Ogura, M. Yakoub, A. Christ, T. Fujiwara, Z. Nikolic
7. Project Administration: K. Ogura
8. Resources: P. Boland, E. Athanasian, J. Healey
9. Software: K. Ogura
10. Supervision: P. Boland, E. Athanasian, J. Healey
11. Validation: K. Ogura, M. Yakoub, A. Christ, T. Fujiwara, Z. Nikolic
12. Visualization: K. Ogura, M. Yakoub, A. Christ, T. Fujiwara, Z. Nikolic
13. Writing – Original Draft: K. Ogura
14. Writing – Review & Editing: J. Healey

INSTITUTIONAL REVIEW BOARD

The study was performed at a single institution, which approved the use of human subjects for this investigation (IRB #16–913). All investigations were conducted in conformity with the ethical principles of research.

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest. No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

distribution-based and anchor-based approaches (the latter utilized receiver operating characteristic analysis).

Results: Using distribution-based methods, the minimal clinically important differences of the DASH questionnaire were 7.4 and 8.3 by half standard deviation and the 90% interval of minimal detectable change, respectively. By anchor-based method (receiver operating characteristic analysis), the minimal clinically important difference was 8.3.

Conclusion: The minimal clinically important difference values calculated by each method validates that the results for upper extremity oncology patients were similar to those reported in other orthopaedic conditions. These results identify the threshold for meaningful improvements in DASH scores in orthopaedic oncology patients and establish the reference to evaluate health-related quality of life and the outcomes of upper extremity oncology surgery. These data should be further refined for disease- and reconstruction-specific analyses.

Level of Evidence: Level III, diagnostic study

Keywords

DASH; minimal clinically important difference; sarcoma; quality-of-life; upper extremity; surveys and questionnaires; patient-reported outcome measures; MODEMS

INTRODUCTION

Due to improved survival of sarcoma patients, there has been increased attention on functional outcomes and health-related quality-of-life (HRQoL) after the treatment of disease. Accordingly, the number of publications on HRQoL after surgery for musculoskeletal tumors has increased.^{7,16,17,26} There have been several measures used for patients with sarcoma in the upper extremity, including the Musculoskeletal Tumor Society (MSTS) rating scale^{11,18,32} and Toronto Extremity Salvage Score.^{1,9,25} Patients have also been evaluated using more generic scales, such as SF-36⁴, or focused scales such as the Patient-Reported Outcomes Measurement Information System (PROMIS) scale^{6,13,27}, and the Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire.³

The DASH questionnaire is the most widely used region-specific measure of disability and symptoms in people with musculoskeletal disorders of the upper limb.³ Definition of the minimal clinically meaningful change varies with the population studied and is disease- or surgical procedure- specific to form homogeneous population.¹⁹ The longitudinal construct validity of the questionnaire scores has been reported to vary considerably from 15–23 (mean of 19) for “much better” or “much worse” after surgery, and 7–14 (mean of 10) for those reporting the arm to be “somewhat better” or “somewhat worse” in selected surgeries ranging from carpal tunnel release to shoulder arthroscopy.¹⁴

This understanding is imperative to interpret DASH scores to assess HRQoL in patients with upper-limb tumors. However, the minimal clinically important difference (MCID), defined as the smallest change a patient can perceive as being meaningful, of DASH in orthopaedic oncology patients has not to our knowledge been reported to date. This study aimed to

determine the MCID of DASH using HRQoL data in patients undergoing surgical treatment for musculoskeletal tumors of the upper extremity.

MATERIALS AND METHODS

This is a retrospective observational study to investigate MCID of the DASH in orthopaedic oncology patients.

Patients and Data Collection

The Musculoskeletal Outcomes Data Evaluation and Management System (MODEMS) instrument, which combines multiple outcomes instruments, including the SF-36 and DASH, was implemented as part of our institutional clinical assessment of all patients. This follows the developed and validated instrument from the American Academy of Orthopaedic Surgeons that was endorsed by the Musculoskeletal Tumor Society after its multispecialty consensus meeting in Tarpon Springs, FL, USA.²⁰

The study was performed at a single institution, which approved the use of human subjects for this investigation (IRB #16–913). All investigations were conducted in conformity with the ethical principles of research.

All patients who underwent surgery and were evaluated for HRQoL using the MODEMS instrument at three months and then every six months after surgery from 1999 to 2005 (n = 960) (Figure 1). At each follow-up survey, information was provided regarding the purposes and methods of the MODEMS instrument.

All patients underwent surgery, and HRQoL outcomes were evaluated using the DASH questionnaire six months following each surgery. When patients received each follow-up questionnaire, an anchor question with a 5-point Likert scale was administered, which inquired about changes in health condition compared with each patients' previous visit. The anchor question included the possible answers: "Much better", "Somewhat better", "About the same", "Somewhat worse", or "Much worse".

Instrument

The DASH questionnaire³ is a region-specific measure of disability and symptoms in people with any or multiple musculoskeletal disorders of the upper limb and allows comparisons across diagnoses of the upper limb. The DASH questionnaire has 30 items rated on a 5-point Likert scale. Scores range from 0 (reflecting no disability) to 100. The items are composed of 1) the degree of difficulty during the preceding week in performing various physical activities because of problems in a shoulder, arm, or hand (21 items); 2) the severity of each of the symptoms of pain, activity-related pain, tingling, weakness, and stiffness (five items); and 3) the problem's effect on social activities, work, and sleep, and its psychological impact (four items). In this study, we did not analyze the modules for work and sports/performing arts, as these models are optional for DASH and not included in MODEMS.

Calculation of MCID

The MCID is used in clinical trials to help readers and patients interpret whether the effect size associated with treatment is a sample large enough to justify the risk, pain, cost, or inconvenience of the treatment^{22,23}. In calculating MCID in this study, two different methodologies were employed: distribution-based and anchor-based approaches.²⁸

The distribution-based approach utilizes the distributional characteristics of the sample. It also determines the observed variation using a standardized metric, such as the standard deviation (SD), effect size, standard error of measurement (SEM), or minimal detectable change (MDC). Calculating MCID using half of SD, SEM, and MDC is widely used because they have been shown to correspond to the MCID across a variety of studies.^{5,24} The rationale for using this approach is that it detects the degree to which change between the baseline and different time points exceeds what would be expected from chance alone. The major disadvantage of all methods that employ the distribution-based approach is that they do not, in themselves, provide a good indication of the importance of the observed change.

The anchor-based approach uses an external criterion, or anchor, to determine what patients or their clinicians consider to be an important improvement/deterioration. Anchor-based methods assess which changes in the measurement instrument correspond with a minimal important change defined by the anchor. The advantage of this method is that the concept of “minimal importance” is explicitly defined and incorporated. The major disadvantage of this method is that global assessment scale may not always be valid. For example, 1) it can be susceptible to recall bias, and 2) it may not be appropriate for all dimensions of multidimensional HRQoL scores.

With these issues in mind, Crosby et al. recommended the combined use of anchor-based and distribution-based methods to take advantage of both an external criterion and a measure of variability, and compensate for disadvantages of each method.⁸ Therefore, we combined both distribution-based and anchor-based approaches to determine the MCIDs of the DASH in the postoperative setting following surgical treatment of upper extremity sarcomas.

With the heterogeneity of orthopaedic oncology conditions in mind, we reported MCIDs of DASH according to the severity of the HRQoL to form more homogeneous groups and to determine the MCID for each group separately, as well as those for our overall patient population. The cutoff value of severity of DASH score was based on the median score and defined as 23.

Distribution-Based Methods

MCIDs were calculated using two different statistical characteristics of the distribution of the scores. First, we determined MCIDs by the half SD method²⁴, where minimal change is considered as the half of SD of change scores. Second, the MDC₉₀ method was used, which sets the MDC at a 90% confidence interval, rather than using SEM. MDC₉₀ was chosen because of its use in previous studies regarding DASH, including a report from the developers of the DASH instrument.^{3,10,12,29} MDC₉₀ is interpreted as 90% of truly stable

patients will demonstrate random variation of less than this magnitude when assessed on multiple occasions.²¹

The SEM and MDC₉₀ were calculated using the following formula, where SD stands for the SD of baseline scores and R stands for the reliability coefficient:

$$\text{SEM} = \text{SD} \times \text{SQRT}(1 - R)$$

$$\text{MDC}_{90} = 1.65 \times \text{SQRT}(2) \times \text{SEM}$$

For the reliability coefficient, 0.96 was used for the DASH score in accordance with previous reports.^{2,3}

Anchor-Based Methods

The anchor-based approach was used to determine cut-off values for MCID based on the answer to the anchor question, which is the standard for assessing the change of patient's condition. Receiver operating characteristic (ROC) curve analysis to examine changes in DASH score between two surveys was performed to compute a discrete value for the MCID by evaluating a threshold difference in the DASH score that yielded the smallest difference between sensitivity and specificity. The ROC-derived MCID was taken to be the change in scores from the baseline with sensitivity and specificity to distinguish between patients who reported their outcome as "About the same" and those who reported their status as "Somewhat better". The discriminative ability of the model was assessed using the area under the ROC curve (AUC).

Statistical Analysis

All statistical analyses were conducted using IBM SPSS version 18.0 (IBM SPSS, Armonk, NY, USA). The scores were reported as the mean values \pm SD. The AUC provides a measure of accuracy, and AUC was interpreted as defined previously¹⁵: 0.90–1.00, excellent; 0.80–0.90, good; 0.60–0.80, fair; and 0.50–0.60, failed. For the required sample size to calculate MCIDs, the validity of the number of cases included in our retrospective database was confirmed. The COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) checklist was followed, which is the standard to study the psychometric property of HRQoL measures. COSMIN considers a sample size of 50 cases as good³¹. Although our cohort meets the COSMIN requirement, in order to reinforce our statistical strategy, we performed power analysis to determine the required sample size in the following conditions using R version 3.6.2 (R Foundation for Statistical Computing, Vienna, Austria): Type I error (Alpha, significance) = 0.05, Type II error (Beta, 1-power) = 0.10, Area under ROC curve = 0.80, Ratio of sample sized in negative/positive groups = 2, one-sided test. The required number of cases was computed as 31 cases (number of positive and negative cases, 10 and 21, respectively), suggesting our sample size was sufficient for calculating MCID.

RESULTS

Seventy-eight patients who completed at least two surveys after surgery and had no missing data were included in the final analysis. The mean duration between surveys was 8.9 months. The clinical and demographic characteristics of the participants are detailed in Table I.

The baseline DASH score, distribution of answers to each anchor question, and change of DASH scores (follow-up score - baseline score) are summarized in Table II. The mean change in DASH score was -6.6, with lower scores signifying clinical improvement. Based on the anchor question, most patients answered either, "About the same" (41.9%) or improved, which included "Much better" or "Somewhat better" (46.0%). The relationship between change of DASH scores and answers to anchor questions is presented using a box-and-whiskers plot (Figure 2), which demonstrates a positive correlation between patient-reported change and change in DASH scores.

MCID of DASH Score

Using distribution-based methods, the MCID of DASH was calculated as 7.4 (7.1 for low baseline score and 7.9 for high baseline score) based on the half SD, and 8.3 (3.3 for low baseline score and 6.1 for high baseline score) based on the MDC₉₀, respectively.

Using the anchor-based method, the MCID of DASH that yielded the smallest difference between sensitivity and specificity, which were calculated using ROC analysis, was estimated as 8.3. On ROC curve analysis, the AUC was 0.75 (95% CI, 0.64–0.86; $P < 0.001$) (Figure 3), which suggested good discriminative ability of the model. When comparing low (DASH score >23 , $N = 39$) and high (DASH score ≤ 23 , $N = 39$) baseline DASH score, ROC-based MCIDs were 3.6 (AUC, 0.69; 95% CI, 0.53–0.85; $P = 0.035$) and 8.8 (AUC, 0.82; 95% CI, 0.67–0.97; $P = 0.001$), respectively. As expected, patients with high baseline DASH score required more change of score to perceive improvement of physical function compared to those with low baseline score.

DISCUSSION

Meaningful threshold change values of outcome measures are essential for understanding changes in each patient's functional status in a concise and comprehensible manner. To that end, MCID was developed to calculate benchmarks for interpreting change in patient-reported outcome measures (PROMs).²⁸ For the DASH, MCID has been calculated in various disease conditions in orthopaedic surgery (Table III).^{3,10,12,29,30} Psychometric evaluation (reliability, validity, responsiveness, or MCID) of PROMs should be tailored to the specific patient population in question, such as patients with different backgrounds or disease conditions. Although the DASH is the most widely used and validated HRQoL measure for upper extremity disorder, it still needs to have its disease specific MCID established. The HRQoL profile of orthopaedic oncology patients is different from those with benign orthopaedic diseases. It is widely known that HRQoL domains exhibit differences by cancer type and anatomy involved. Yet, musculoskeletal oncology patients in general, and upper extremity patients in particular, have not been analyzed. It is also

reasonable that orthopaedic oncology patients have different HRQoL profiles compared to those of general orthopaedic patients with unaltered life expectancy and more consistent anatomic and functional deficits.^{22,23} Therefore, we reported MCIDs of DASH specific to upper extremity orthopaedic oncology patients.

The MCID data presented will aid in the interpretation of clinical data regarding different upper extremity procedures. The MCID constitutes a threshold for outcome scores where a patient would consider a given change in score to be meaningful and worthwhile. It is critical to study MCIDs because they can be used for sample size calculation in clinical trials. MCIDs are also used to facilitate the establishment of treatment recommendations. We can investigate the effectiveness of interventions (i.e., physical therapy or drug therapy) by comparing the change of scores in patients with and without interventions by evaluating those that exceed MCID value or not. Therefore, the MCID data will also allow clinicians to better interpret changes in patients' DASH scores over time.

There is no agreement as to the best approach for calculation the MCID. Crosby et al. recommended combined use of distribution-based and anchor based methods.⁸ Of the several types of distribution-based methods, we used half SD and MDC₉₀-based calculations because they are among the most universally applied distribution-based approaches. For anchor-based methods, we used the global rating of change scale to quantify patient improvement/deterioration over time.

At the time of assessment, patients were asked to independently rate the overall change in their lower extremity condition. While the anchor-based method correlated each patient's perceived clinical change with changes in HRQoL scores, MCID values derived from the global rating of change scale may suffer from recall bias. In this study, the calculated distribution-based and anchor-based MCID values for the DASH questionnaires were similar (around eight points). Our results were also similar to MCID values previously reported in other orthopaedic conditions, allowing us to conclude that the MCIDs of the DASH calculated in this study in upper extremity orthopaedic oncology patients are reasonable.^{3,10,12,29,30} However, one should be aware that the MCID values can be within a range. Caution should be exercised when interpreting and using published MCID values at the individual level, especially when different anatomic sites and diseases are being studied. Also, there is a clear need for improvement and standardization of MCID methodology in the DASH questionnaire, as well as other PROMs, to allow finer discrimination and evaluation of clinical responsiveness.

The MCID constitutes a threshold for outcome scores where a patient would consider a given change in score to be meaningful. It is critical to study MCIDs because they can be used for sample size calculation in clinical trials. MCIDs are also used to facilitate the establishment of treatment recommendations. We can investigate the effectiveness of interventions (i.e. physical therapy or drug therapy) by comparing the change of scores in patients with and without interventions by evaluating those exceed MCID value or not. Since both anchor-based and distribution-based MCID estimates of the DASH in patients with musculoskeletal tumors of the upper extremity were quite similar, we have confidence in the estimates made, which were about 8 points. This suggests interventions improving

DASH by less than that amount are unlikely to be perceived by patients as clinically meaningful. Therefore, those interventions may not justify exposing patients to risk, cost, or inconvenience. When applying novel interventions for orthopaedic oncology patients, including drugs or rehabilitation regimens, considering MCIDs will be vital to evaluate the clinical and potential statistical significance of the change. Also, the MCID-based approach will be helpful to identify subgroups of patients who could benefit from a specific intervention, even if the overall population did not benefit.

Limitations

This study has several limitations. First, we used several methods to calculate the MCID of DASH. Currently, there is no consensus regarding the optimal way to estimate the MCID. Although the half SD method is the most frequently used (approximately 90% of reports on MCID)^{5,24}, it is affected by the sample studied; it will vary depending on the sample size and distribution of results. The MDC_{90} based approach may be more effective; however, it requires knowledge of the measure's reliability.^{2,3} Alternatively, the global rating of change scale, which was used as an anchor question, may not always be valid with strong correlation within all the DASH subscales, including physical function, severity of pain or stiffness, or problem associated with social activities, work, and sleep. Due to the heterogeneity of diagnoses and treatments, which are endemic in studies of patients undergoing musculoskeletal tumor surgery, patients with variable levels of function were included in this study increases the ability to generalize their results, but may limit them when finer discrimination is needed. In our sub-analysis, we also found baseline score (severity) affected MCID values of DASH. This heterogeneity reinforces the need for further evaluation of MCID in uniform patient populations based on disease- and reconstruction-specific analyses.

CONCLUSIONS

Since both distribution-based and anchor-based MCID estimates of the DASH in patients with musculoskeletal tumors were so similar, there is confidence in the estimates made, which were about 8 points. This suggests that interventions improving DASH by less than that amount are unlikely to be perceived by patients as clinically important. Therefore, those interventions may not justify exposing patients to risk, cost, or inconvenience. When applying novel interventions for orthopaedic oncology patients such as operations, implants, drugs, or rehabilitation regimens, considering MCIDs will be vital to evaluate. The MCID-based approach will also be helpful to identify subgroups of patients who could benefit from a specific intervention, even if the overall population did not benefit. Our findings provide benchmark values for MCID in musculoskeletal oncology, which serves as a reference for future studies of HRQoL in upper extremity musculoskeletal oncology patients using the DASH questionnaire. Further refinement of the analysis should be disease- and reconstruction-specific.

ACKNOWLEDGMENTS

FUNDING

This study was funded in part through the NIH/NCI Cancer Center Support Grant, P30 CA008748.

REFERENCES

1. Akiyama T, Uehara K, Ogura K, Shinoda Y, Iwata S, Saita K et al. Cross-cultural adaptation and validation of the Japanese version of the Toronto Extremity Salvage Score (TESS) for patients with malignant musculoskeletal tumors in the upper extremities. *J Orthop Sci* 2017;22:127–132. 10.1016/j.jos.2016.09.012 [PubMed: 27743633]
2. Angst F, Goldhahn J, Drerup S, Kolling C, Aeschlimann A, Simmen B et al. Responsiveness of five outcome measurement instruments in total elbow arthroplasty. *Arthritis Care Res (Hoboken)* 2012;64:1749–1755. 10.1002/acr.21744 [PubMed: 22674818]
3. 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;14:128–146. [PubMed: 11382253]
4. Brazier JE, Harper R, Jones NM, O’Cathain A, Thomas KJ, Usherwood T et al. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *Bmj* 1992;305:160–164. 10.1136/bmj.305.6846.160 [PubMed: 1285753]
5. Brigden A, Parslow RM, Gaunt D, Collin SM, Jones A, Crawley E. Defining the minimally clinically important difference of the SF-36 physical function subscale for paediatric CFS/ME: triangulation using three different methods. *Health Qual Life Outcomes* 2018;16:202. 10.1186/s12955-018-1028-2 [PubMed: 30340599]
6. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care* 2007;45:S3–s11. 10.1097/01.mlr.0000258615.42478.55
7. Coens C, van der Graaf WT, Blay JY, Chawla SP, Judson I, Sanfilippo R et al. Health-related quality-of-life results from PALETTE: A randomized, double-blind, phase 3 trial of pazopanib versus placebo in patients with soft tissue sarcoma whose disease has progressed during or after prior chemotherapy—a European Organization for research and treatment of cancer soft tissue and bone sarcoma group global network study (EORTC 62072). *Cancer* 2015;121:2933–2941. 10.1002/cncr.29426 [PubMed: 26033286]
8. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 2003;56:395–407. [PubMed: 12812812]
9. Davis AM, Wright JG, Williams JI, Bombardier C, Griffin A, Bell RS. Development of a measure of physical function for patients with bone and soft tissue sarcoma. *Qual Life Res* 1996;5:508–516. [PubMed: 8973131]
10. 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;17:1257–1267. 10.1007/s11136-008-9409-3 [PubMed: 18958582]
11. Enneking WF, Dunham W, Gebhardt MC, Malawar M, Pritchard DJ. A system for the functional evaluation of reconstructive procedures after surgical treatment of tumors of the musculoskeletal system. *Clin Orthop Relat Res* 1993;241–246. [PubMed: 8425352]
12. Franchignoni F, Vercelli S, Giordano A, Sartorio F, Bravini E, Ferriero G. Minimal clinically important difference of the disabilities of the arm, shoulder and hand outcome measure (DASH) and its shortened version (QuickDASH). *J Orthop Sports Phys Ther* 2014;44:30–39. 10.2519/jospt.2014.4893 [PubMed: 24175606]
13. Garcia SF, Cella D, Clauser SB, Flynn KE, Lad T, Lai J et al. Standardizing patient-reported outcomes assessment in cancer clinical trials: a patient-reported outcomes measurement information system initiative. *J Clin Oncol* 2007;25:5106–5112. 10.1200/jco.2007.12.2341 [PubMed: 17991929]
14. Gummesson C, Atroshi I, Ekdahl C. The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: longitudinal construct validity and measuring self-rated health change after surgery. *BMC Musculoskelet Disord* 2003;4:11. 10.1186/1471-2474-4-11 [PubMed: 12809562]
15. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29–36. 10.1148/radiology.143.1.7063747 [PubMed: 7063747]

16. Holzer LA, Huyer N, Friesenbichler J, Leithner A. Body image, self-esteem, and quality of life in patients with primary malignant bone tumors. *Arch Orthop Trauma Surg*2019:[Epub ahead of print]. 10.1007/s00402-019-03205-8
17. Hudgens S, Forsythe A, Kontoudis I, D'Adamo D, Bird A, Gelderblom H. Evaluation of Quality of Life at Progression in Patients with Soft Tissue Sarcoma. *Sarcoma*2017;2017:2372135. 10.1155/2017/2372135 [PubMed: 28512389]
18. Iwata S, Uehara K, Ogura K, Akiyama T, Shinoda Y, Yonemoto Tet al.Reliability and Validity of a Japanese-language and Culturally Adapted Version of the Musculoskeletal Tumor Society Scoring System for the Lower Extremity. *Clin Orthop Relat Res*2016;474:2044–2052. 10.1007/s11999-016-4880-6 [PubMed: 27169541]
19. Jayadevappa R, Cook R, Chhatre S. Minimal important difference to infer changes in health-related quality of life-a systematic review. *J Clin Epidemiol*2017;89:188–198. 10.1016/j.jclinepi.2017.06.009 [PubMed: 28676426]
20. Johanson NA, Liang MH, Daltroy L, Rudicel S, Richmond J. American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments. Reliability, validity, and sensitivity to change. *J Bone Joint Surg Am*2004;86:902–909. 10.2106/00004623-200405000-00003 [PubMed: 15118030]
21. Kennedy DM, Stratford PW, Wessel J, Gollish JD, Penney D. Assessing stability and change of four performance measures: a longitudinal study evaluating outcome following total hip and knee arthroplasty. *BMC Musculoskelet Disord*2005;6:3. 10.1186/1471-2474-6-3 [PubMed: 15679884]
22. Leopold SS. Editorial: Importance of Validating the Scores We Use to Assess Patients with Musculoskeletal Tumors. *Clin Orthop Relat Res*2019;477:669–671. 10.1097/corr.0000000000000631 [PubMed: 30624325]
23. Leopold SS, Porcher R. Editorial: The Minimum Clinically Important Difference-The Least We Can Do. *Clin Orthop Relat Res*2017;475:929–932. 10.1007/s11999-017-5253-5 [PubMed: 28124297]
24. Norman GR, Sloan JA, Wywrich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*2003;41:582–592. 10.1097/01.Mlr.0000062554.74615.4c [PubMed: 12719681]
25. Ogura K, Uehara K, Akiyama T, Iwata S, Shinoda Y, Kobayashi Eet al.Cross-cultural adaptation and validation of the Japanese version of the Toronto Extremity Salvage Score (TESS) for patients with malignant musculoskeletal tumors in the lower extremities. *J Orthop Sci*2015;20:1098–1105. 10.1007/s00776-015-0767-8 [PubMed: 26307208]
26. Postma A, Kingma A, De Ruiter JH, Schraffordt Koops H, Veth RP, Goeken LNet al.Quality of life in bone tumor patients comparing limb salvage and amputation of the lower extremity. *J Surg Oncol*1992;51:47–51. [PubMed: 1518295]
27. Reeve BB, Hays RD, Bjorner JB, Cook KF, Crane PK, Teresi JAet al.Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care*2007;45:S22–31. 10.1097/01.mlr.0000250483.85507.04 [PubMed: 17443115]
28. 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;61:102–109. 10.1016/j.jclinepi.2007.03.012 [PubMed: 18177782]
29. Schmitt JS, Di Fabio RP. Reliable change and minimum important difference (MID) proportions facilitated group responsiveness comparisons using individual threshold criteria. *J Clin Epidemiol*2004;57:1008–1018. 10.1016/j.jclinepi.2004.02.007 [PubMed: 15528051]
30. Sorensen AA, Howard D, Tan WH, Ketchersid J, Calfee RP. Minimal clinically important differences of 3 patient-rated outcomes instruments. *J Hand Surg Am*2013;38:641–649. 10.1016/j.jhsa.2012.12.032 [PubMed: 23481405]
31. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res*2012;21:651–657. 10.1007/s11136-011-9960-1 [PubMed: 21732199]

32. Uehara K, Ogura K, Akiyama T, Shinoda Y, Iwata S, Kobayashi E et al. Reliability and Validity of the Musculoskeletal Tumor Society Scoring System for the Upper Extremity in Japanese Patients. *Clin Orthop Relat Res* 2017;475:2253–2259. 10.1007/s11999-017-5390-x [PubMed: 28560530]

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

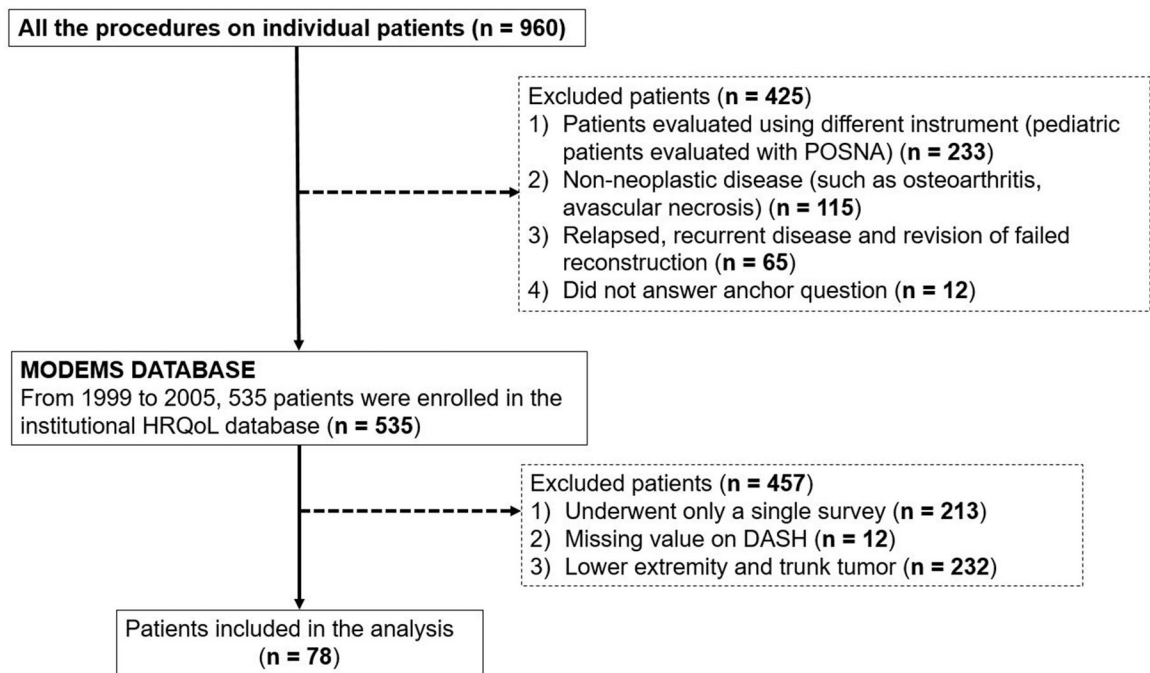


Figure 1.

Flow diagram of patient inclusion and exclusion criteria. POSNA = Pediatric Orthopaedic Society of North America; HRQoL = health-related quality-of-life; DASH = Disabilities of the Arm, Shoulder and Hand.

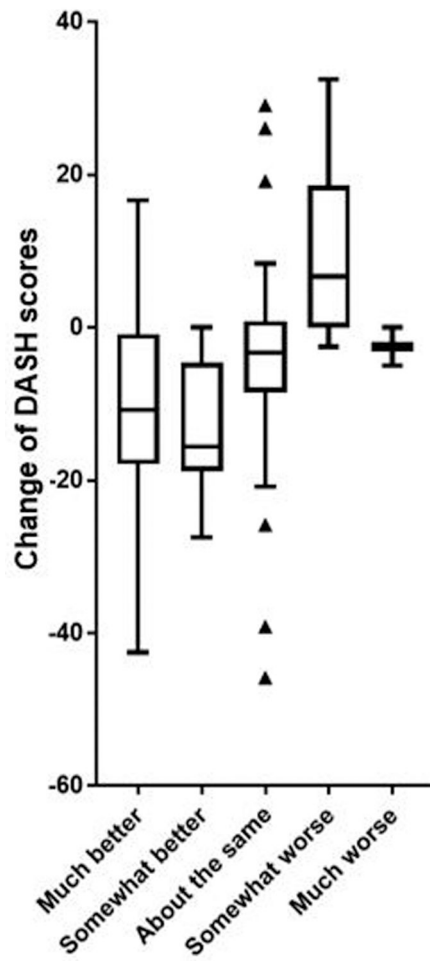


Figure 2. Box-and-whiskers plots to demonstrate relationship between change of DASH score and answers to anchor questions. DASH = Disabilities of Arm, Shoulder and Hand.

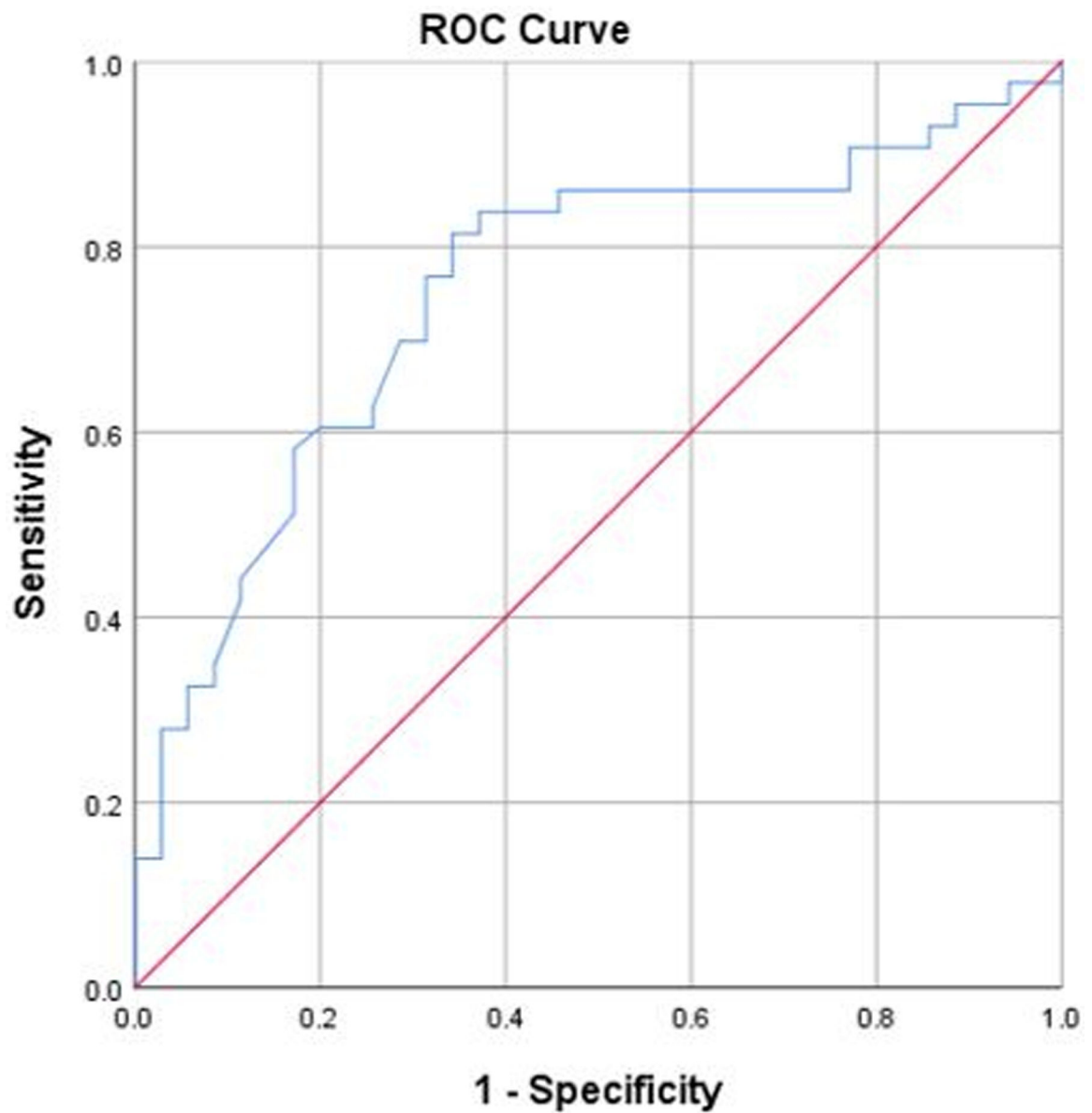


Figure 3.

The ROC curve plots the true positive rate (sensitivity) against the false-positive rate (1 - specificity) for Disabilities of The Arm, Shoulder and Hand questionnaire scores. ROC = receiver operating characteristic.

TABLE I.

Patient demographics

Characteristic	Number of patients (%)
Overall	78 (100)
Age, years; mean [SD ^a]	48.4 (range, 18–87) [17.0]
Sex	
Male	37 (47.4)
Female	41 (52.6)
Histologic diagnosis	
Chondrosarcoma	18 (23.1)
Synovial sarcoma	7 (9.0)
Osteosarcoma	6 (7.7)
Undifferentiated pleomorphic sarcoma	6 (7.7)
Myxofibrosarcoma	5 (6.4)
Liposarcoma	4 (5.1)
Fibrosarcoma	4 (5.1)
Desmoid tumor	4 (5.1)
Ewing's sarcoma	2 (2.6)
Leiomyosarcoma	2 (2.6)
Skin cancer	2 (2.6)
Benign lesion	1 (1.3)
Other soft tissue sarcoma	5 (6.4)
Cancer metastasis	12 (15.4)
Type of surgery	
Wide resection	42 (53.8)
Wide resection + endoprosthetic reconstruction	9 (11.5)
Wide resection + allograft reconstruction	6 (7.7)
Curettage +/- arthroplasty/internal fixation	18 (23.1)
ORIF ^b	1 (1.3)
Amputation	2 (2.6)

^a-SD = standard deviation;

^b-ORIF = open reduction and internal fixation.

TABLE II.Details of DASH^a score and anchor questions

Characteristic	Number of patients (%)
Baseline survey; mean [SD ^b]	24.9 [18.1]
Follow-up survey	
Anchor question	
“Much better”	19 (25.7)
“Somewhat better”	15 (20.3)
“About the same”	31 (41.9)
“Somewhat worse”	8 (10.8)
“Much worse”	1 (1.4)
Change of the DASH score; mean [SD]	-6.6 [14.8]

^aDASH-Disabilities of Arm, Shoulder and Hand;^bSD- standard deviation.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE III.

Comparison of MCID^a and DASH^b among orthopaedic conditions

Report	Author	Number of cases	Subjects	MCID Calculation	MCID
<i>J Hand Ther</i> , 2001	Beaton DE et al. ³	200	Patients with upper-limb musculoskeletal disorders	MDC ^c 90	10.7
<i>J Clin Epidemiol</i> , 2004	Schmitt JS et al. ²⁹	154	Adults with musculoskeletal upper extremity problems	Anchor	12.6
				MDC90	12.6
<i>Qual Life Res</i> , 2008	Dawson J et al. ¹⁰	104	Patients undergoing elbow surgery	MDC90	9.3
				Anchor	10.3
<i>J Hand Surg Am</i> , 2013	Sorensen A et al. ³⁰	102	Patients undergoing non-operative treatment for tendonitis, arthritis, or nerve compression from the forearm to the hand.	Anchor	10.0
<i>J Orthop Sports Phys Ther</i> , 2013	Franchignoni F et al. ¹²	255	Patients with upper limb musculoskeletal disorders	Anchor	10.8
				MDC90	10.8
Current series	Ogura K et al.	78	Patients undergoing surgery for upper-limb tumor	Anchor	8.3
				MDC90	8.3
				Half SD ^d	7.4

^aMCID = minimal clinically important difference;^bDASH = Disabilities of the Arm, Shoulder and Hand;^cMDC = minimal detectable change;^dSD = standard deviation.