



# HHS Public Access

Author manuscript

*J Orthop Res.* Author manuscript; available in PMC 2022 October 01.

Published in final edited form as:

*J Orthop Res.* 2021 October ; 39(10): 2116–2123. doi:10.1002/jor.24944.

## Minimal Clinically Important Differences in SF-36 Global Score: Current Value In Orthopedic Oncology

Koichi Ogura, MD, PhD<sup>1</sup>, Meredith K. Bartelstein, MD<sup>1</sup>, Mohamed Yakoub, MD<sup>1</sup>, Zarko Nikolic, MD<sup>1</sup>, Patrick J. Boland, MD<sup>1</sup>, John H. Healey, MD, FACS<sup>1</sup>

<sup>1</sup>Department of Surgery, Orthopaedic Service, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA.

### Abstract

The SF-36 is widely used to evaluate the health-related quality of life of patients with musculoskeletal tumors. Instead of typical methods, calculating the SF-36 Global Score has recently become an increasingly common reporting approach. However, numerical changes lack clear clinical relevance. The minimal clinically important difference is useful for interpreting changes in functional scores by defining the smallest change patients may perceive as clinically meaningful. The aim of this study is to determine the minimal clinically important difference of the SF-36 Global Score in orthopaedic oncology patients, which has not been reported to date. Three-hundred ten patients who underwent surgery and completed two surveys during postoperative follow-up were reviewed. We used the two most common methods to calculate the SF-36 Global Score: 1) anchor-based methods and receiver operating characteristic analysis based on one-half of the standard deviation of change score and standard error of measurement at baseline and; 2) distribution-based methods. Using anchor-based methods, the minimal clinically important differences of SF-36 Global Scores #1 and #2 were 2.7 (area under the curve [AUC] = 0.85) and 2.5 (AUC = 0.79) for improvement, and -1.5 (AUC = 0.81) and -0.6 (AUC = 0.83) for deterioration, respectively. Using distribution-based methods, the minimal clinically important differences of SF-36 Global Scores #1 and #2 were 4.1 and 4.4 by half standard deviation, and 4.1 and 4.5 by the standard error of measurement, respectively. Our findings provide benchmark values, which can serve as a reference for future studies in musculoskeletal tumor patients using the SF-36 Global Score as a single measure for health-related quality of life.

**Trial Registration:** Retrospectively registered.

---

Corresponding Author: Dr. John H. Healey, MD, FACS, 1275 York Avenue, New York, NY, 10065, USA, P: 212-639-7611, F: 212-717-3573, healeyj@mskcc.org.

Authors' contributions

K.O. analyzed and interpreted the patient data regarding the study and was a major contributor in writing the manuscript. M.B., M.Y., and Z.N. analyzed and interpreted the patient data regarding the study and helped write and edit the manuscript. P.B. and J.H. provided supervision to the project, analyzed and interpreted patient data regarding the study, and was a major contributor in writing the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This study was approved by the Institutional Review Board/Ethics Committee of Memorial Sloan Kettering Cancer Center, where the ethics committee waived written informed consent for participants under Institutional Review Board Protocol number 16-913.

Competing interests

The authors declare that they have no competing interests.

## Keywords

sarcoma; SF-36 Global Score; health-related quality of life; minimal clinically important difference; anchor-based method; distribution-based method

---

## BACKGROUND

Patient-derived outcome assessments are essential determinants of health-related quality of life (HRQoL) in oncology and help inform treatment decisions for future musculoskeletal oncology patients. There has been a burgeoning of publications regarding HRQoL.<sup>1</sup> More generic instruments, such as SF-36 or Patient Reported Outcomes Measurement Information System (PROMIS), are validated internationally for general orthopaedic conditions.<sup>2; 3</sup> However, these have not been evaluated critically in musculoskeletal oncology, although they have been used to assess correlations with the six subscales of the Musculoskeletal Tumor Society (MSTS) score.<sup>4</sup>

Statistical significance is typically calculated to study how scores change over time. However, the clinical significance of changes in these scores is not clearly understood. The minimal clinically important difference (MCID), defined as the smallest change that a patient can perceive as being meaningful, has been determined in several orthopaedic conditions, but not musculoskeletal oncology.<sup>5-8</sup>

The most common use of SF-36 is based on two component scores (mental component summary [MCS] and physical component summary [PCS]) from eight subscales. However, a single underlying construct of HRQoL in SF-36 was indicated which would correspond to an SF-36 Global Score indicating global health of the patients. Since then, SF-36 Global Scores are being used increasingly to assess global health of the patients in published studies due to simplicity.<sup>9-11</sup> The aim of this study was to determine the MCID for SF-36 Global Scores in orthopaedic oncology patients treated with surgery for a wide range of musculoskeletal tumors, which has not been reported to date.

## METHODS

### Patients and data collection

This study's Level of Evidence was Level III, diagnostic study. Each author certifies that the study institution approved the human protocol for this investigation, and all investigations were conducted in conformity with the ethical principles of research. The analysis reported in this study was approved by Memorial Sloan Kettering Cancer Center's institutional review board protocol number 16-913.

The Musculoskeletal Outcomes Data Evaluation and Management System (MODEMS) instrument, which combines multiple outcomes instruments, including the SF-36, was implemented as part of our institutional clinical assessment for all patients. This was in accordance with the developed and validated instrument from the American Academy of Orthopaedic Surgeons and endorsed in 1994 by the Musculoskeletal Tumor Society after its multispecialty consensus meeting in Tarpon Springs, FL, USA.<sup>12</sup>

All patients who underwent surgery and evaluated for HRQoL using the MODEMS instrument at three months and then every six months after surgery were reviewed from 1999 to 2005. At each follow-up appointment, patients were presented with a survey where they were asked the following anchor question, “Compared to when you last completed this questionnaire, is your musculoskeletal condition, ‘much better,’ ‘somewhat better,’ ‘about the same,’ ‘somewhat worse,’ or ‘much worse’?”. Of 535 patients enrolled in the MODEMS database, we excluded patients with no follow-up data (patients underwent only a single survey,  $n = 213$ ), or those with insufficient HRQoL data (missing value on SF-36 PCS or MCS or anchor question,  $n = 12$ ) (Figure 1).

## Instrument

MODEMS was used, which incorporates components from the SF-36 (v1, four-week recall), along with other postsurgical functional and HRQoL parameters. The SF-36 measures of health on eight multi-item dimensions covering functional status, well-being, and overall evaluation of health (i.e. physical functioning [PF], role-physical [RP], bodily pain [BP], general health [GH], vitality [VT], social functioning [SF], role-emotional [RE], and mental health [MH]).<sup>13</sup> Patients were asked to rate their responses using 2-, 3-, 5-, or 6-point polytomous response options. For each dimension, item scores are calculated, summed, and transformed to a scale ranging from zero (worst health) to 100 (best health). These had a mean score of 50 and standard deviation (SD) of 10, in accordance to guidelines presented by John Ware, SF-36/ SF-12 manuals, and interpretation guides.<sup>14; 15</sup>

We used the two most common methods for estimating SF-36 Global Score: the arithmetic average for eight subscales (SF-36 Global Score #1) and component summary (SF-36 Global Score #2) using the following formulas, respectively<sup>16</sup>:

$$SF-36\ Global\ Score\ \#1 = (PF + RP + BP + GH + VT + SF + RE + MH)/8$$

$$SF-36\ Global\ Score\ \#2 = (PCS + MCS)/2$$

## MCID Calculation

MCID helps with sample size calculation and interpreting results in clinical trials.<sup>17</sup> Anchor-based and distribution-based methods are the established approaches to calculate MCID. Anchor-based methods also establish independent criteria that patients and doctors believe to be important indicators of improvement or worsening of each patient’s condition. The criteria identify the minimal important variation from the anchor. This method has the advantage of explicitly identifying what constitutes the difference that is of, “minimal importance.” In distinction, the statistical distribution in the population uses a standardized metric to determine the observed variation, which may be the standard error of the mean (SEM), SD, or effect size. Many studies have shown that one half of the SEM or SD corresponds to the MCID.<sup>18; 19</sup> This strategy identifies the extent of a change from baseline values beyond chance and fails to establish whether the observed change is actually important.

Some recommendations have been proposed regarding the best method to apply for determining robust and reliable MCIDs. For example, many clinicians believe that anchor-based calculations of the MCID, which define, “clinically important” in relation to changes identified as important by the patients themselves, are more relevant to clinical practice than distribution-based methods.<sup>20–22</sup> Although this approach can seem subjective, asking patients about changes in their health status to determine their individual perceptions is critical. However, others criticize the anchor-based method because of the impact of potential response shift effect due to the longitudinal design used which may bias the results.<sup>23</sup>

With these issues in mind, Crosby et al. recommended the combined use of anchor-based and distribution-based methods to take advantage of an external criterion, a measure of variability, and compensate for disadvantages of each method.<sup>20</sup> Based on the review article on cancer studies related to determination of MCIDs<sup>24</sup>, >80% of the studies examined used both distribution- and anchor-based methods, suggesting the importance of cross-referencing the MCIDs calculated by two different methods to verify the validity of each value. Therefore, we combined both anchor- and distribution-based methods to determine the MCIDs of the SF-36 Global Score in the postoperative setting following surgical treatment of musculoskeletal tumors.

### Anchor-based methods

Anchor-based methods were used to determine cut-off values for MCIDs based on the answer to each anchor question, which is the gold standard for assessing the change of each patient’s condition.<sup>25</sup> Receiver operating characteristic (ROC) curve analysis for a change in SF-36 Global Score between two surveys were performed to compute a discrete value for the MCID by evaluating a threshold delta in the SF-36 Global Score that yielded the smallest difference between sensitivity and specificity. The ROC-derived MCIDs were 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” for improvement and, “Somewhat worse” for deterioration. The discriminative ability of the model was assessed using the area under the ROC curve (AUC).

### Distribution-based methods

MCIDs were calculated using two different statistical characteristics of the distribution of the scores at the baseline. First, the MCIDs by the half SD method were calculated,<sup>19</sup> where minimal change is considered as the half of SD of change scores. Second, the minimum amount of change potentially detectable using the SEM was estimated. The SEM was calculated using the following formula, where SD is for the baseline scores, and R stands for the reliability coefficient:

$$SEM = SD \times \sqrt{1 - R}$$

For the reliability coefficient, 0.90 was used for the SF-36 Global Score in accordance with previous reports.<sup>5; 26</sup>

## 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 was interpreted as defined previously<sup>27</sup>: 0.90–1.00, excellent; 0.80–0.90, good; 0.60–0.80, fair; and 0.50–0.60, failed.

## RESULTS

The mean duration between the surveys was nine months. The clinical and demographic characteristics of the participants are detailed in Table 1.

The baseline SF-36 Global Scores, distribution of answers to each anchor question, and change of SF-36 scores (follow-up score - baseline score) are summarized in Table 2. Based on the anchor question, most patients answered either, “About the same” (39.7%) or improved, noted in “Much better” or, “Somewhat better” responses (40.6%). The relationship between changes of SF-36 Global Scores and answers to anchor questions is presented using a box-and-whiskers plot. This demonstrated a positive correlation between changes in both SF-36 Global Scores #1 and #2 and patient-reported change in disease condition (Figure 2).

### MCID of SF-36 Global Scores

Table 3 shows the MCID of SF-36 Global Scores. Using the anchor-based method, the discrete MCID values that yielded the smallest difference between sensitivity and specificity are shown and suggested good discriminative ability of the model (Figure 3).<sup>27</sup> As expected, patients required an increased change in score to perceive improvement than deterioration (Table 3).

## DISCUSSION

As survival improves for orthopaedic oncology patients, increased interest has developed in patient-reported outcomes (PROs), particularly for function and HRQoL.<sup>28–30</sup> The SF-36 is a generic measure of HRQoL and is widely used in healthcare to evaluate patient’s self-perceived health.<sup>31; 32</sup> Interpreting outcome measures presents a challenge, as statistical significance does not always translate into clinical significance. Therefore, meaningful interpretation of these PROs requires an understanding of the MCID required to interpret a change in status. The SF-36 Global score works effectively to characterize outcome for musculoskeletal oncology patients.<sup>16</sup>

Anagnostopoulos et al. used a structural equation model analysis to investigate the latent structure of SF-36 in a population-based sample of adults<sup>33</sup>. Three orders of latent structure of SF-36 were identified: the first-order factors were eight SF-36 subscales. These displayed three second-order factors as PCS, MCS, and “well-being” based on GH and VT domain scores. They also showed a third-order factor, which would correspond to an SF-36 Global Score, indicating that all SF-36 responses address a single underlying construct of HRQoL<sup>33</sup>

In the scoping review by Lins et al.<sup>16</sup>, only one of the reports discussed involved a study in oncology.<sup>34</sup> Cookson et al. stated that analyses of SF-36 dimensionality denote cross-loadings from the eight domains and also between PCS and MCS.<sup>34</sup> These findings would suggest the existence of uni-dimensionality that would support the calculation of a single index of HRQoL. However, they noted that there has not been full validation of it. Our study of the MCID of the SF-36 Global score is an important element that helps to validate its clinical utility. These results justify the use of the SF-36 Global Score as a simple single measure of HRQoL. Furthermore, current trends in data reporting support the use of the SF-36 Global Score as a simple, single measure of HRQoL. Other instruments recommend the calculation of a single index of quality of life, such as the Pediatric Quality of Life Inventory, EuroQol-5D, and the 15-D.<sup>16</sup> Given this trend of using the SF-36 Global Score as a simple, single measure of HRQoL, there is a compelling need for guiding metrics to interpret the clinical significance of changes of SF-36 Global Score. These are essential for the SF-36 Global Score to be used in meaningful contexts in clinical care.

Based on the user's manual for the SF-36 health survey, a distributional approach applied to a representative United States general population sample produced the following estimates of MCID values for each SF-36 score: 2 points for RP, 4 points for RE, 3 points for other 6 subscales, and 3 points for PCS and MCS.<sup>35</sup> Based on these estimates, the estimated MCIDs of SF-36 for each of the calculated total scores (#1, the average of MCIDs for each of the eight subscales, and #2, average of MCIDs for the PCS and MCS) would be estimated to be three points each.<sup>35</sup> However, there is no prior data available regarding the MCID of SF-36 Global Score. The MCID data for improvement presented in the current study seems to be valid based on the estimates of MCID values provided in the SF-36 health survey user's manual. The lower MCID values for deterioration suggested that the SF-36 is more responsive to deterioration than to improvement, a phenomenon also reported for other HRQoL measures.<sup>36–38</sup>

Although the SF-36 was not originally designed to examine economic evaluation, SF-6D was developed for use in economic evaluation studies, which is a single index ranging from 0.0 (worst health state) to 1.0 (best health state). It can be used in the assessment of the quality adjusted life years and the cost-effectiveness of various healthcare interventions, including EQ-5D.<sup>39; 40</sup> The SF-6D focuses on seven of the eight subscales covered by the SF-36: PF, SF, BP, MH, VT, and RP/RE. Only the GH domain is not included. The SF-6D is widely validated in several studies, including MCID ranging from 0.026–0.041.<sup>41–44</sup> However, it is designed for economic evaluation studies and not for measuring health-related quality of life.

Several limitations of this study should be noted. First, the global rating of change scale, which we used as an anchor question, may not always be valid for every subscale in SF-36. The use of multiple anchor questions would be ideal, considering SF-36 contains multiple dimensions of HRQoL. However, no previous literature has validated the MCID of SF-36 using such sophisticated methodology. We also confirmed moderate to strong correlation between anchor question and change in the scores to verify the anchor we used is valid for both Global Scores (Spearman's correlation coefficient, 0.54 and 0.63 for Global Scores #1 and #2, respectively). Second, MCID values derived from the global rating of change

scale using anchor-based approaches may suffer from intrinsic weakness, such as subjective retrospective judgment of change. The validity of the scale used for the global assessment may or may not be accurate; it is also vulnerable to recall bias, and not equally appropriate for each dimension of HRQoL. Third, due to the heterogeneity of diagnoses together with the diverse range of treatments (a common problem in studies of musculoskeletal tumor surgery), patients with variable levels of function were included in this study. Widely-varied operations can be evaluated by the same instruments. However, the outcomes are different.<sup>45</sup> This heterogeneity reinforces the need for further evaluation of MCID in a more uniform patient population. Finally, the MCIDs presented cannot necessarily be applied to patients with other disease conditions; the MCID can be impacted by population characteristics, such as disease type and severity.<sup>46</sup> Physicians who use these metrics need to carefully interpret them, especially when evaluating patients with non-orthopaedic oncologic disease conditions.

The MCID values were calculated for the SF-36 Global Scores based on both anchor- and distribution-based approaches. The consistency of our results across different approaches suggests that a change in score of 2.5 to 4.5 for improvement and -0.6 to -1.5 for deterioration is meaningful for patients with musculoskeletal tumors. This suggested SF-36 is more responsive to deterioration than to improvement, a phenomenon also reported for other measures. Our findings provide benchmark values for MCID of SF-36 Global Score, which can serve as a reference for future studies of HRQoL in musculoskeletal tumor patients using the SF-36 Global Score as a single measure for HRQoL.

## Acknowledgments

The authors thank Ms. Jessica Massler, MSW, for her editorial assistance.

## Funding

This research was funded in part by the NIH/NCI Cancer Center Support Grant, P30 CA008748, The Perlman Limb Preservation Fund, and the Jake's Reindeer Run Foundation. All funding bodies had no role in the study's design nor in the data collection, analysis, and interpretation. The funding bodies were not involved in the writing of the manuscript.

## Availability of data and materials

The data supporting the findings of this study are available upon request, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the study institution.

## REFERENCES

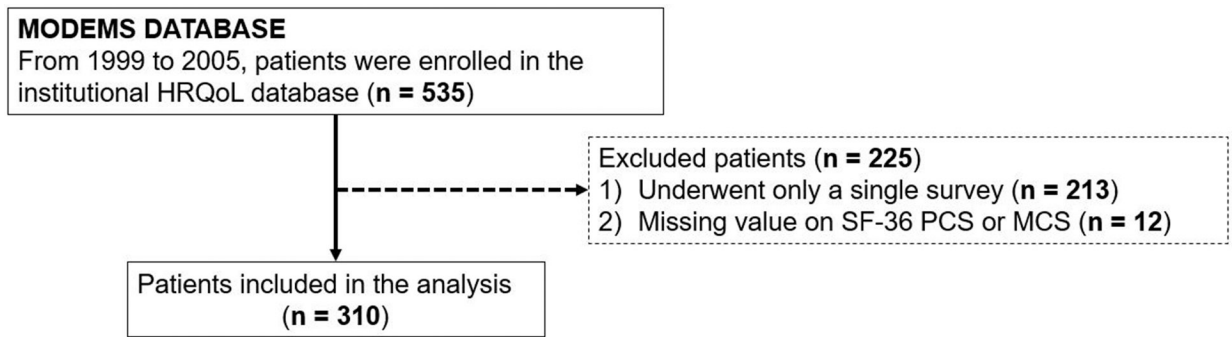
1. Coens C, van der Graaf WT, Blay JY, et al. 2015. Health-related quality-of-life results from PALETTE: A randomized, double-blind, phase 3 trial of pazopanib versus placebo in patients with soft tissue sarcoma whose disease has progressed during or after prior chemotherapy—a European Organization for research and treatment of cancer soft tissue and bone sarcoma group global network study (EORTC 62072). *Cancer* 121:2933–2941. [PubMed: 26033286]
2. Laucis NC, Hays RD, Bhattacharyya T. 2015. Scoring the SF-36 in Orthopaedics: A Brief Guide. *J Bone Joint Surg Am* 97:1628–1634. [PubMed: 26446970]

3. Picavet HS, Hoeymans N. 2004. Health related quality of life in multiple musculoskeletal diseases: SF-36 and EQ-5D in the DMC3 study. *Ann Rheum Dis* 63:723–729. [PubMed: 15140781]
4. Uehara K, Ogura K, Akiyama T, et al. 2017. Reliability and Validity of the Musculoskeletal Tumor Society Scoring System for the Upper Extremity in Japanese Patients. *Clin Orthop Relat Res* 475:2253–2259. [PubMed: 28560530]
5. Badhiwala JH, Witiw CD, Nassiri F, et al. 2018. Minimum Clinically Important Difference in SF-36 Scores for Use in Degenerative Cervical Myelopathy. *Spine* 43:E1260–E1266. [PubMed: 29652783]
6. Copay AG, Glassman SD, Subach BR, et al. 2008. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. *Spine J* 8:968–974. [PubMed: 18201937]
7. Kosinski M, Zhao SZ, Dedhiya S, et al. 2000. Determining minimally important changes in generic and disease-specific health-related quality of life questionnaires in clinical trials of rheumatoid arthritis. *Arthritis Rheum* 43:1478–1487. [PubMed: 10902749]
8. Yuksel S, Ayhan S, Nabiye V, et al. 2019. Minimum clinically important difference of the health-related quality of life scales in adult spinal deformity calculated by latent class analysis: is it appropriate to use the same values for surgical and nonsurgical patients? *Spine J* 19:71–78. [PubMed: 30010046]
9. Pellegrini MJ, Schiff AP, Adams SB Jr., et al. 2015. Conversion of Tibiotalar Arthrodesis to Total Ankle Arthroplasty. *J Bone Joint Surg Am* 97:2004–2013. [PubMed: 26677234]
10. Queen RM, Sparling TL, Butler RJ, et al. 2014. Patient-Reported Outcomes, Function, and Gait Mechanics After Fixed and Mobile-Bearing Total Ankle Replacement. *J Bone Joint Surg Am* 96:987–993. [PubMed: 24951733]
11. Tidermark J, Bergstrom G, Svensson O, et al. 2003. Responsiveness of the EuroQol (EQ 5-D) and the SF-36 in elderly patients with displaced femoral neck fractures. *Qual Life Res* 12:1069–1079. [PubMed: 14651424]
12. Johanson NA, Liang MH, Daltroy L, et al. 2004. American Academy of Orthopaedic Surgeons lower limb outcomes assessment instruments. Reliability, validity, and sensitivity to change. *J Bone Joint Surg Am* 86:902–909. [PubMed: 15118030]
13. Brazier JE, Harper R, Jones NM, et al. 1992. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *Bmj* 305:160–164. [PubMed: 1285753]
14. Ware JE, Bjoner JB, Turner-Bowker DM, Gandek B, Maruish ME. 2008. User's manual for the SF-36 v2 Health Survey. Philadelphia, PA: Springer;
15. Ware JE. 2001. SF-36 physical & mental health summary scales: a manual for users of version 1. Philadelphia, PA: Springer;
16. Lins L, Carvalho FM. 2016. SF-36 total score as a single measure of health-related quality of life: Scoping review. *SAGE Open Med* 4:2050312116671725. [PubMed: 27757230]
17. Copay AG, Subach BR, Glassman SD, et al. 2007. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 7:541–546. [PubMed: 17448732]
18. Brigden A, Parslow RM, Gaunt D, et al. 2018. 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* 16:202. [PubMed: 30340599]
19. Norman GR, Sloan JA, Wyrwich KW. 2003. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 41:582–592. [PubMed: 12719681]
20. Crosby RD, Kolotkin RL, Williams GR. 2003. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 56:395–407. [PubMed: 12812812]
21. Maltenfort M, Diaz-Ledezma C. 2017. Statistics In Brief: Minimum Clinically Important Difference-Availability of Reliable Estimates. *Clin Orthop Relat Res* 475:933–946. [PubMed: 28050812]
22. Rai SK, Yazdany J, Fortin PR, et al. 2015. Approaches for estimating minimal clinically important differences in systemic lupus erythematosus. *Arthritis Res Ther* 17:143. [PubMed: 26036334]



23. Hamidou Z, Dabakuyo TS, Bonnetain F. 2011. Impact of response shift on longitudinal quality-of-life assessment in cancer clinical trials. *Expert Rev Pharmacoecon Outcomes Res* 11:549–559. [PubMed: 21958100]
24. Ousmen A, Touraine C, Deliu N, et al. 2018. Distribution- and anchor-based methods to determine the minimally important difference on patient-reported outcome questionnaires in oncology: a structured review. *Health Qual Life Outcomes* 16:228. [PubMed: 30537955]
25. Leopold SS, Porcher R. 2017. Editorial: The Minimum Clinically Important Difference-The Least We Can Do. *Clin Orthop Relat Res* 475:929–932. [PubMed: 28124297]
26. Ware JE Jr. 2000. SF-36 health survey update. *Spine (Phila Pa 1976)* 25:3130–3139. [PubMed: 11124729]
27. Hanley JA, McNeil BJ. 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143:29–36. [PubMed: 7063747]
28. Lemieux J, Brundage MD, Parulekar WR, et al. 2018. Quality of Life From Canadian Cancer Trials Group MA.17R: A Randomized Trial of Extending Adjuvant Letrozole to 10 Years. *J Clin Oncol* 36:563–571. [PubMed: 29328860]
29. Ranft A, Seidel C, Hoffmann C, et al. 2017. Quality of Survivorship in a Rare Disease: Clinicofunctional Outcome and Physical Activity in an Observational Cohort Study of 618 Long-Term Survivors of Ewing Sarcoma. *J Clin Oncol* 35:1704–1712. [PubMed: 28355116]
30. van Wulfften Palthe ODR, Janssen SJ, Wunder JS, et al. 2017. What questionnaires to use when measuring quality of life in sacral tumor patients: the updated sacral tumor survey. *Spine J* 17:636–644. [PubMed: 27856381]
31. Bekkering WP, van Egmond-van Dam JC, Bramer JAM, et al. 2017. Quality of life after bone sarcoma surgery around the knee: A long-term follow-up study. *Eur J Cancer Care (Engl)* 26.
32. Rivard JD, Puloski SS, Temple WJ, et al. 2015. Quality of life, functional outcomes, and wound complications in patients with soft tissue sarcomas treated with preoperative chemoradiation: a prospective study. *Ann Surg Oncol* 22:2869–2875. [PubMed: 25783679]
33. Anagnostopoulos F, Niakas D, Pappa E. 2005. Construct validation of the Greek SF-36 Health Survey. *Qual Life Res* 14:1959–1965. [PubMed: 16155784]
34. Cookson MS, Dutta SC, Chang SS, et al. 2003. Health related quality of life in patients treated with radical cystectomy and urinary diversion for urothelial carcinoma of the bladder: development and validation of a new disease specific questionnaire. *J Urol* 170:1926–1930. [PubMed: 14532809]
35. Maruish ME. 2011. User's manual for the SF-36v2 Health Survey (3rd ed.). Lincoln, RI: QualityMetric Incorporated;
36. Campbell H, Rivero-Arias O, Johnston K, et al. 2006. Responsiveness of objective, disease-specific, and generic outcome measures in patients with chronic low back pain: an assessment for improving, stable, and deteriorating patients. *Spine (Phila Pa 1976)* 31:815–822. [PubMed: 16582856]
37. Hagg O, Fritzell P, Nordwall A. 2003. The clinical importance of changes in outcome scores after treatment for chronic low back pain. *Eur Spine J* 12:12–20. [PubMed: 12592542]
38. Mannion AF, Porchet F, Kleinstuck FS, et al. 2009. The quality of spine surgery from the patient's perspective: part 2. Minimal clinically important difference for improvement and deterioration as measured with the Core Outcome Measures Index. *Eur Spine J* 18 Suppl 3:374–379. [PubMed: 19296136]
39. Brazier J, Roberts J, Deverill M. 2002. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 21:271–292. [PubMed: 11939242]
40. Brazier J, Usherwood T, Harper R, et al. 1998. Deriving a preference-based single index from the UK SF-36 Health Survey. *J Clin Epidemiol* 51:1115–1128. [PubMed: 9817129]
41. Harvie HS, Honeycutt AA, Neuwahl SJ, et al. 2019. Responsiveness and minimally important difference of SF-6D and EQ-5D utility scores for the treatment of pelvic organ prolapse. *Am J Obstet Gynecol* 220:265.e261–265.e211. [PubMed: 30471259]
42. Walters SJ, Brazier JE. 2003. What is the relationship between the minimally important difference and health state utility values? The case of the SF-6D. *Health Qual Life Outcomes* 1:4. [PubMed: 12737635]

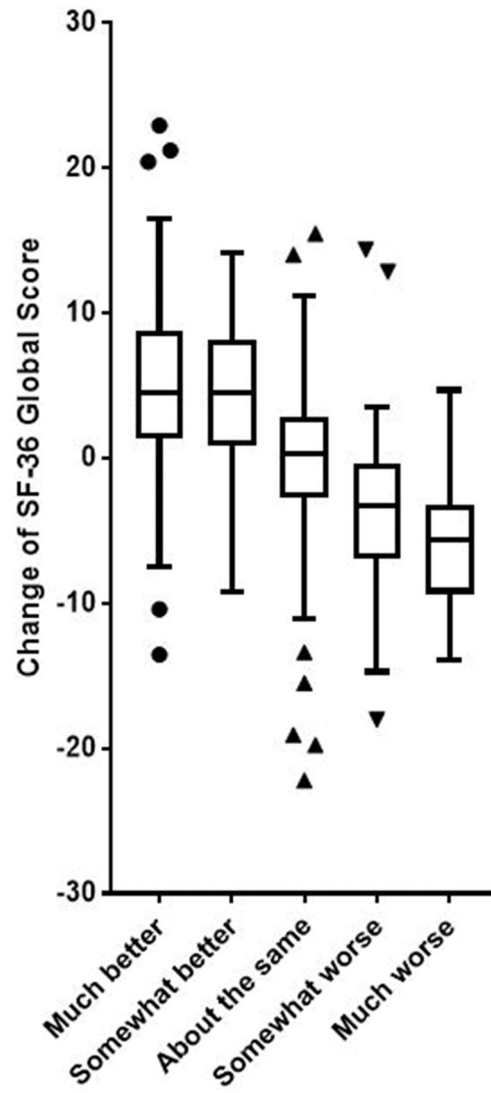
43. Walters SJ, Brazier JE. 2005. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. *Qual Life Res* 14:1523–1532. [PubMed: 16110932]
44. Drummond M 2001. Introducing economic and quality of life measurements into clinical studies. *Ann Med* 33:344–349. [PubMed: 11491193]
45. Malek F, Somerson JS, Mitchel S, et al. 2012. Does limb-salvage surgery offer patients better quality of life and functional capacity than amputation? *Clin Orthop Relat Res* 470:2000–2006. [PubMed: 22302658]
46. Jayadevappa R, Cook R, Chhatre S. 2017. Minimal important difference to infer changes in health-related quality of life-a systematic review. *J Clin Epidemiol* 89:188–198. [PubMed: 28676426]

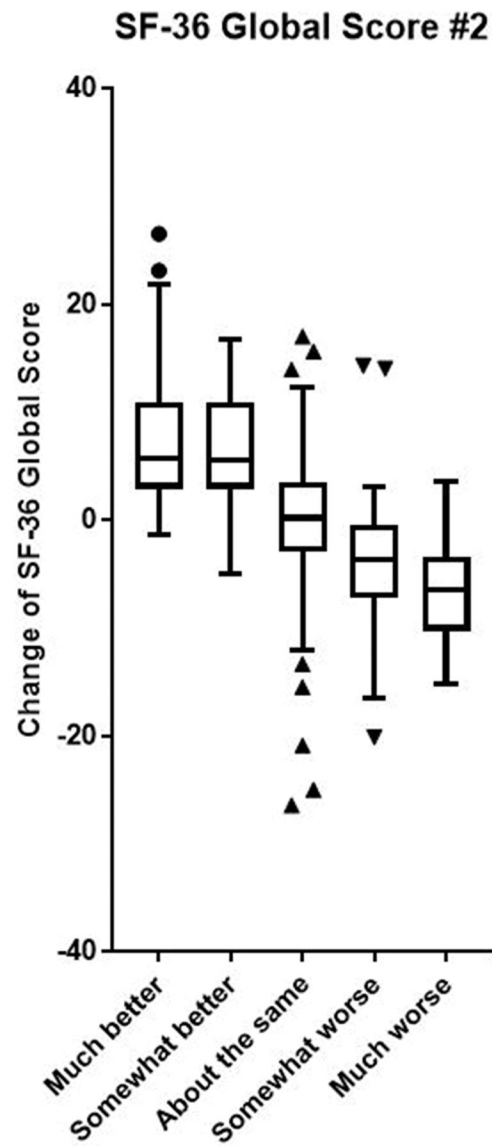


**Figure 1.**  
Flow diagram showing patient eligibility criteria for analysis.

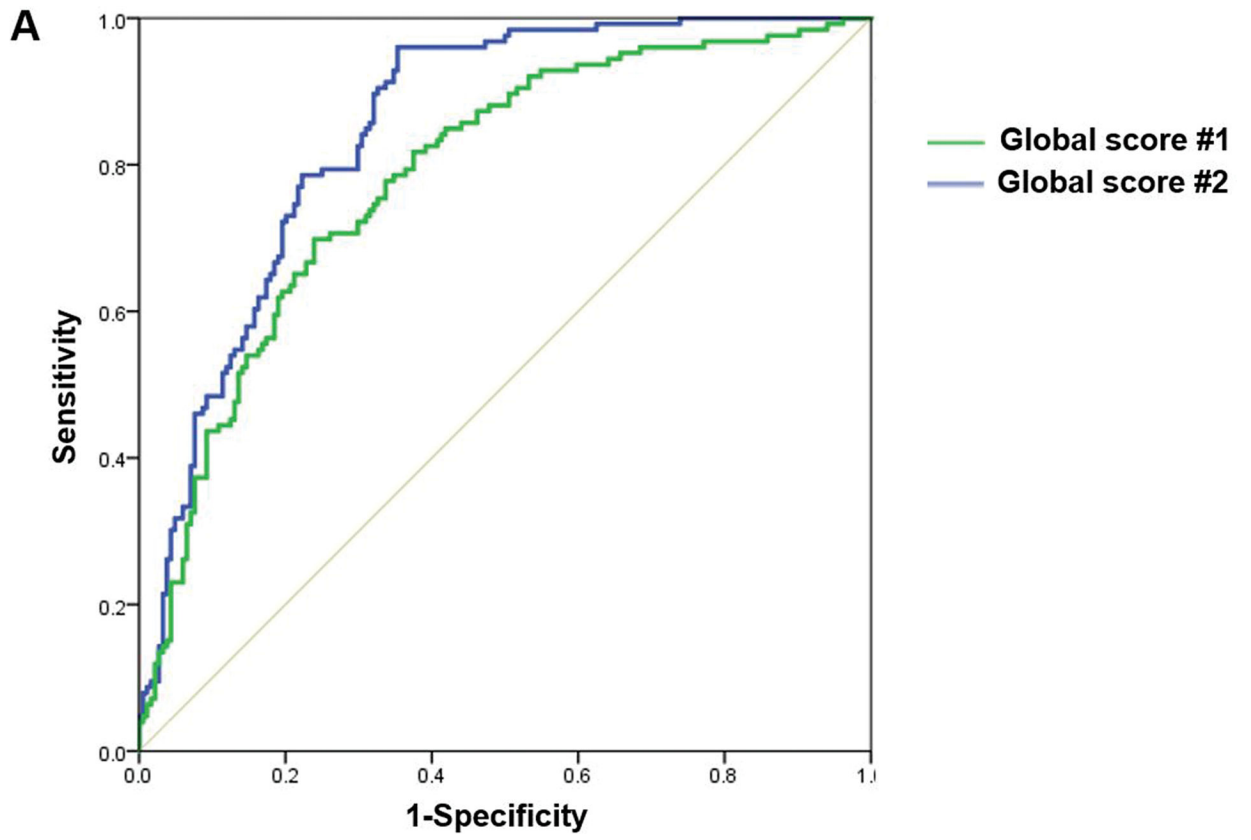
**A**

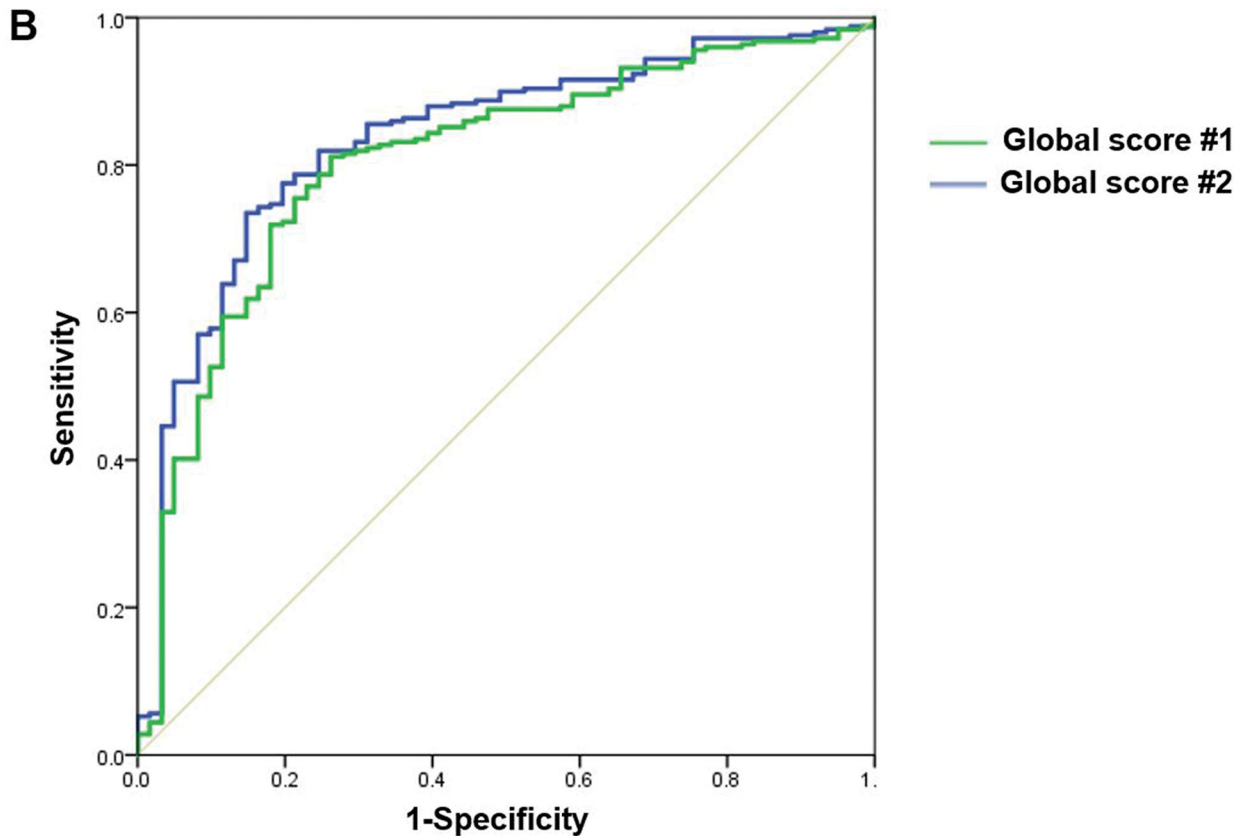
**SF-36 Global Score #1**



**B****Figure 2.**

Box-and-whiskers plots demonstrate the relationship between change of SF-36 Global Scores #1 (A) and #2 (B) and also displays answers to each anchor question.





**Figure 3.**

The ROC curves for improvement (A) and deterioration (B) plots the true positive rate (sensitivity) against the false-positive rate (1 - specificity) for SF-36 Global Scores #1 (green line), and #2 (blue line).

**TABLE 1.**

Patient demographics

Characteristic	Number of patients (%)
Overall	310 (100)
Age, years; mean [standard deviation]	54.1 (range, 14–87) [16.7]
Sex	
Male	157 (50.6)
Female	153 (49.4)
<b>Histologic diagnosis</b>	
Chondrosarcoma	38 (12.3)
Undifferentiated pleomorphic sarcoma	26 (8.4)
Benign lesion	24 (7.7)
Osteosarcoma	23 (7.4)
Chordoma	14 (4.5)
Synovial sarcoma	11 (3.5)
Ewing's sarcoma	9 (2.9)
Liposarcoma	7 (2.3)
Leiomyosarcoma	7 (2.3)
Fibrosarcoma	6 (1.9)
Myxofibrosarcoma	6 (1.9)
Desmoid tumor	5 (1.6)
Skin cancer	5 (1.6)
Other soft tissue sarcoma	11 (3.5)
Cancer metastasis	118 (38.1)
<b>Type of surgery</b>	
Wide resection	129 (41.6)
Wide resection + endoprosthetic reconstruction	73 (23.5)
Wide resection + allograft reconstruction	17 (5.5)
Curettage +/- arthroplasty/internal fixation	36 (11.6)
Open reduction and internal fixation	12 (3.9)



Characteristic	Number of patients (%)
Amputation/rotationplasty	10 (3.2)
Others*	33 (10.6)

\* Spinal decompression, surgery for nonunion, etc.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**TABLE 2.**

Details of SF-36 scores and anchor question

Characteristic	Number of patients (%)
<b>Baseline survey</b>	
<b>SF-36</b>	
Global Score #1; mean [SD] <sup>a</sup>	44.0 [9.3]
Global Score #2; mean [SD]	44.3 [9.9]
<b>Follow-up survey</b>	
<b>Anchor question</b>	
"Much better"	55 (17.7)
"Somewhat better"	71 (22.9)
"About the same"	123 (39.7)
"Somewhat worse"	46 (14.8)
"Much worse"	15 (4.8)
<b>Change of the SF-36 scores</b>	
Global Score #1; mean [SD]	+1.9 [7.6]
Global Score #2; mean [SD]	+1.2 [6.7]

<sup>a</sup>SD, standard deviation.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE 3.

MCID<sup>a</sup> of SF-36 Global Scores

Method of calculation	SF-36 Global Score	
	#1	#2
<b>Anchor-based method, ROC<sup>b</sup></b>		
Improvement	2.7 (AUC <sup>c</sup> , 0.85)	2.5 (AUC, 0.79)
Deterioration	-1.5 (AUC, 0.81)	-0.6 (AUC, 0.83)
<b>Distribution-based method</b>		
Half SD <sup>d</sup>	4.1	4.4
SEM <sup>e</sup>	4.1	4.5

<sup>a</sup>-MCID, minimal clinically important difference;<sup>b</sup>- receiver operating characteristic;<sup>c</sup>-AUC, area under the curve;<sup>d</sup>-SD, standard deviation;<sup>e</sup>-SEM, standard error of measurement.