Validation of the Hip Disability and Knee Injury and Osteoarthritis Outcome Score (HOOS, KOOS) pain and function subscales for use in Total Hip (THR) and Total Knee Replacement (TKR) clinical trials.

Susan M Goodman[1], **Bella Mehta**[1], **Lisa A Mandl**[1], **Jackie Szymonifka**[1,5], **Jackie Finik**[1,6], **Mark Figgie**[1], **Iris Navarro-Millán**[1], **Mathias Bostrom**[1], **Michael Parks**[1], **Douglas Padgett**[1], **Alexander McLawhorn**[1], **Vinicius Antao**[1], **Adolph Yates**[2], **Bryan Springer**[3], **Steven Lyman**[1], **Jasvinder A Singh**[4]

[1]Hospital for Special Surgery, 535 East 70th Street, New York, NY 10021

[2]University of Pittsburgh, 5200 Centre Avenue Suite 415, Pittsburgh, PA 15232

[3]OrthoCarolina Hip and Knee Center, 2001 Vale Avenue; Suite 200, Charlotte, NC 28207

[4]The University of Alabama at Birmingham Medical Center, 2000 6th Ave South, Birmingham, AL 35294

[5]NYU Langone Medical Center, 540 1st Avenue, New York, NY 10016

[6]Memorial Sloan Kettering Cancer Center, 485 Lexington Avenue, New York, NY 10017, USA

Abstract

Background: THR/TKR studies do not uniformly measure patient centered domains, pain and function. We aim to validate existing measures of pain and function within subscales of standard instruments to facilitate measurement.

Methods: We evaluated baseline and 2-year pain and function for THR, TKR using HOOS/KOOS scores, with primary unilateral TKR (4,796) and THR (4,801). Construct validity was assessed by correlating HOOS/KOOS pain and activities of daily living (ADL), function quality of life (QOL) and satisfaction using Spearman correlation coefficients. Patient relevant thresholds for change in pain and function were anchored to improvement in quality of life; Mean Clinically Important Difference (MCID) corresponded to "a little improvement" and a Really Important Difference (RID) to a "moderate improvement". Pain and ADL function scores were compared by quartiles using Kruskal-Wallis.

Results: 2 year HOOS/KOOS pain and ADL function correlated with health related QOL (KOOS pain and SF12PCS ρ=0.54; function ρ=0.63). Comparing QOL by pain and function quartiles, the highest levels of pain relief and function were associated with the most improved QOL. MCID for pain was estimated at 20, and the RID 29; MCID for function 14, and the RID 23. The measures were responsive to change with large effect sizes (1.8).

Corresponding Author: **Dr. Susan M Goodman** – Hospital for Special Surgery goodmans@hss.edu P: 212-606-1163 F: 212-472-6567.

Conclusions: We confirm HOOS/KOOS pain and ADL function subscales are valid measures of critical patient centered domains after THR/TKR, and achievable thresholds anchored to improved quality of life. Cost-free availability and brevity makes them feasible, to be used in a core measurement set in TJR trials.

Background:

Total joint replacements are widely utilized as therapy for end-stage arthritis, relieving pain and restoring function for the majority of patients. However, 20–30% of TKR patients, and 10% of THR report persistent pain and/or are dissatisfied with their surgical outcome (1, 2). Ideally, data from trials, registries, and outcomes studies could be systematically pooled to facilitate understanding risk factors for poor outcomes; to date, however, there has not been a universally agreed core outcome set to facilitate effective data pooling to make comparisons between trials. To this end, the Outcomes Measurement in Rheumatology (OMERACT) Total Joint Replacement (TJR) Special Interest Group (SIG) was formed in 2008 (3, 4). This international collaboration of stakeholders, including orthopedists, physical therapists, rheumatologists, patients, and methodologists, has met and combined the results of systematic literature reviews and Delphi consensus surveys to prioritize a core domain set for all TJR outcome studies. The core domain set endorsed by OMERACT in 2016 included pain, function, patient satisfaction, revision, adverse events, and death (5). The TJR core domain set was vetted by stakeholders, including both patients and orthopedic surgeons, who agreed that these are critical outcomes (5–7). The candidate outcome measures for the core domains are the pain and activities of daily living (ADL) function subscales of the Hip disability and Osteoarthritis Outcome Score (HOOS) and Knee injury and Osteoarthritis Outcome Scores. Both subscales are part of existing validated larger surveys. The ADL domain is an accepted measure of function relevant to TKR, and includes the items in the Western Ontario McMasters (WOMAC) function domain (8–10). Validating the subscales in a TJR cohort offers the opportunity to develop a core measurement set specifically for TJR that includes the outcomes most important to patients that will be suitable for use in TJR clinical trials, registries, and outcomes reports.

The purpose of this study was to evaluate whether pain and ADL function subscales of the HOOS/KOOS possess construct validity, responsiveness to change and meaningful thresholds for change after total hip (THR) and total knee replacement (TKR) respectively, using a large sample of people who underwent these procedures. The complete HOOS (40 questions)/KOOS (42 questions) questionnaires measure 5 dimensions of lower extremity pain, joint specific symptoms, function in activities of daily living (ADL), as well as sports and recreation and quality of life in people with hip/knee disease and are currently in wide use after THR and TKR, but most validation studies include the sports and recreation domain that is not relevant to a majority of the patients undergoing TJR (8, 10, 11). Both instruments were meticulously designed with items generated in an iterative process beginning with literature review and included input from stakeholder groups comprised of

patients, orthopedists and physical therapists. Both instruments have undergone extensive psychometric testing. These measures are free to use, and have previously been shown to be highly reliable, with excellent internal consistency (Cronbach's alpha coefficient of 0.82 to 0.98) in samples of people undergoing THR and TKR (12, 13).

While HOOS/KOOS instruments are in wide use, the relationship of the HOOS/KOOS subscales to other important patient reported outcomes such as change in quality of life (QOL) or satisfaction is less well established, and clinically meaningful thresholds to change have not been clearly linked to the patient important measure of improved quality of life, in a large representative THR/TKR cohort. Moreover, the complete surveys would place a lengthy responder burden if included with the other core set items. We hypothesized that the subscales of HOOS/KOOS pain and ADL function will correlate with changes in patient-reported QOL and satisfaction, demonstrating convergent construct validity. We also hypothesized that HOOS/KOOS will not correlate with mental/emotional QOL two years after THR and TKR, demonstrating divergent validity for these subscale measures. In addition, we hypothesized that HOOS/KOOS pain and ADL function subscales will be more sensitive to change compared with patient reported QOL measured with the Short Form 12 (SF-12) and EuroQol-5D (EQ-5D), and satisfaction two years after THR and TKR as anchors. If shown to be valid and sensitive to change in patient populations with THR and TKR, and stable up to 2 years after THR/TKR, HOOS/KOOS pain and ADL function subscales can constitute a TJR core measurement set that can harmonize outcomes reporting of TJR clinical trials.

Methods:

Patients and Data Sources

We identified all patients undergoing unilateral THR or TKR between May 1, 2007 and February 25, 2011, and retrospectively analyzed prospectively collected data in a Hospital for Special Surgery (HSS) hospital-based total joint replacement (TJR) registry. All patients undergoing arthroplasty were approached to participate in the registry, and data were collected at baseline prior to the surgery and at 2-years post-TJR. Baseline data collected on all patients include age, sex, ethnicity, race, and education level (some college or above or no college). Administrative data included the Charlson-Deyo comorbidity index, which is based on International Classification of Diseases, ninth revision, clinical modification (ICD-9-CM) codes, is a validated measure of co-morbidity commonly used in clinical research (14, 15). At our institution, surgery is almost exclusively elective, and the scores rarely exceed 3, although the total score possible is 26. We also included the American Society of Anesthesiology (ASA) score, a ranking used to quantify surgical risk that ranges from 0–6, with a score of 0 indicating excellent health and a score of 6 indicating an organ transplant donor (16). We followed Strengthening of Reporting of Observational Studies (STROBE) guidelines for reporting the results of this retrospective analysis of a prospective cohort study (17).

The study cohort included patients undergoing primary unilateral arthroplasty who provided baseline HOOS/KOOS and 2-year data including the HSS satisfaction scale, quality of life, and/or the HOOS/KOOS scores at 2 years, without other exclusions. For this study,

approximately 64% of the Joint registry patients met our criteria and were included (Appendix Figure 1a, b). Overall differences in characteristics between those in our study cohort with specific 2-year variables and those without were calculated, and each separate analysis was performed with those cases with complete data for that variable.

All patients included in the registry provided informed consent. The study was approved by the HSS Institutional Review Board. A summary of the measures included in this analysis are in Appendix Table 1.

Pain and Function Outcome Measures: Construct validity

Pain and function were measured using the HOOS/KOOS validated instruments specific for lower extremity pain or ADL function that have demonstrated responsiveness to change after THR and TKR. Each transformed HOOS/KOOS subscale score range is 1–100, with 100 being the best score (8, 9, 16). HOOS/KOOS function is assessed with two separate subscales: (1) activities of daily living (ADL); and (2) sports and recreation. The ADL function subscale is being evaluated in this study as the sports and recreation domain has been found to not be relevant to many joint replacement patients (18–20). The KOOS pain subscale has 9 items, and the KOOS ADL function includes 17 items; HOOS pain contains 10 items and HOOS ADL function 16 items. Both surveys have a similar Quality of Life (QOL) subscale, comprised of 4 items.

Change in the Quality of life (QOL) was queried with both the HOOS/KOOS QOL subscale and with a separate single item with 6 response choices which asked how TJR changed the quality of life: "More than I ever dreamed possible", "Great improvement", "Moderate improvement", "A little improvement", "No improvement at all", or "My quality of life is worse" scored 1–6, higher scores indicating the worst QOL (22). Satisfaction with TJR outcome was measured using 4 primary questions, each rated on a 5-point Likert scale. Each question was weighted equally and a satisfaction summary score was calculated (range 0–100, higher scores corresponding to greater satisfaction). *Pain relief, Improving ability to do housework or yardwork, Improving ability to do recreational activities*, and *Overall satisfaction with surgery* gave response choices of "Very satisfied", "Somewhat satisfied", "Neither satisfied or dissatisfied", "Somewhat dissatisfied", or "Very dissatisfied" (11). The SF-12 is a general health measure that consists of 12 items(21, 22). The SF-12 physical component scale (PCS) and the mental component scale (MCS) are population norm-based summary scale scores and are scored 1–100. The scoring is designed to represent a population mean of 50 with a standard deviation of 10. Higher scores on the SF-12 PCS indicate better physical health status, and lower scores on the MCS correlate with depression (21, 23). EQ-5D is a utility measure of health-related quality of life and records the patient's self-rated health with a 5 level survey and a vertical visual analogue scale and is used as a quantitative measure of health outcome, scored from 0 to 100, higher score indicating "the best health possible" and the lowest score indicating a state worse than death(22, 24, 25). Patient expectations of TJR outcomes were measured using the HSS Total Hip and Total Knee Replacement Expectations Surveys, validated instruments, which specifically question a patient's expectations prior to THR and TKR in areas specific to arthroplasty(26, 27).

Convergent validity of the HOOS/KOOS pain and function subscales was calculated by correlating the 2-year scores with the two measures of a similar construct that we hypothesized to be moderately correlated, in our case, by assessing the Spearman's correlation coefficient with changes in the QOL and patient satisfaction, SF 12 PCS and EuroQUOL. at 2-years after THR or TKR. We examined divergent validity through the association of the SF-12 MCS, a measure of emotional health, that would not be associated with HOOS pain and function either at baseline or at 2 years, as emotional health would not be as strongly associated with pain and function as physical health measured as the SF-12 PCS (28). While poor 2-year KOOS pain and function has been reported in association with poor baseline SF-12 MCS, we hypothesized that this association would be weak (28). We additionally assessed the association of pain relief and improved function at 2 years with improved quality of life and satisfaction, and compared the scores by quartiles.

Pain and Function Outcome Measures: Sensitivity to change and Meaningful Thresholds for change

Sensitivity to change was assessed by calculating Effect size (ES) and standardized response mean (SRM), an index of the effect size, to differentiate clinical change from statistical change. The ES quantifies the size of the difference in outcomes between groups and is independent of the size of the sample. The ES was calculated as mean change in scores (HOOS/KOOS pain, HOOS/KOOS function, SF-12 pain subscale) from baseline to 2-years divided by the baseline SD. SRM was calculated as the mean change in the scores divided by the standard deviation (SD) of the changed scores. According to Cohen's h, an ES of 0.20–0.49 represents a small change, 0.50–0.79 a medium change and   0.80 a large change (29, 30).

We used the HSS Satisfaction QOL responses ("More than I ever dreamed possible" through "My quality of life is worse") to identify clinically meaningful thresholds for the improvement/worsening from baseline to 2-years in HOOS/KOOS pain and function subscales. We defined the minimally clinically important difference for improvement (**MCID Improvement)** by the response "a little improvement" on the Quality of Life question, and response "moderate improvement" on the same question indicated the Really Important Difference for improvement (**RID Improvement)**, which we anchored to these patient important domains. We limited this analysis to people with some (MCID) and large (RID), improvements.

Pain and Function Outcome Measures: Floor and Ceiling effects

The presence of floor and ceiling effects were determined by reviewing the variance plots for homogeneity, to examine the distribution of each domain. Median and IQR were used to summarize Pain and Function scores for each quartile of reported PROMs domains.

Statistical Analysis:

Descriptive statistics were used for comparisons between groups. Correlation was analyzed with Spearman coefficients, categorized according to Cohen's h; 0.20–0.49 represents a small change, 0.50–0.79 a medium change and   0.80 a large change (29, 30). Scores were

compared by quartiles using the Kruskal-Wallis test. All statistical analysis was completed using the SASS software, version 9.4.

Results

Characteristics of the cohort

We included 10,775 subjects in the HSS register who underwent primary unilateral hip or knee replacement, had baseline data, and completed any 2-year HOOS/KOOS pain or function assessments, satisfaction, or quality of life responses. For the included patients who provided some 2-year data, complete data was available on 97% of patients. There were no significant differences at baseline for those in the study cohort with complete 2-year survey responses and those who did not regarding race or ethnicity, comorbidities, baseline HOOS and KOOS function, HSS Expectations Score, or SF-12 PCS or MCS scores. However, a large number of older patients and those with more pain at baseline did not return the surveys, but the numbers were small (Table 1).

Mean age was 67 years for the patients undergoing TJR, 59.4% were women, and 91% were white. Overall outcomes for HOOS/KOOS pain and function at 2-years were excellent with > 40 point mean improvement for pain and function in all demographic groups, with similar improvements at 2-years for HOOS/KOOS function (Appendix Table 2 a,b). For patients undergoing TKR, 77% reported a great improvement or "more than I ever dreamed possible" for their quality of life. For patients undergoing THR, 89% reported a great improvement in their quality of life or "more than I ever dreamed possible" (Table 2). Overall satisfaction was high; 78% of TKR and 91% of THR were very satisfied with pain relief, and only 6.5% of TKR and 2.5 % of THR were somewhat or very dissatisfied. Distribution of quality of life responses for TKR and THR patients in our cohort can be seen in Figure 2.

Convergent and Divergent Validity: HOOS/KOOS pain, function, and QOL scores at 2-years were moderately correlated with the scores for EQ5D and SF-12 PCS, with little or negligible correlation with baseline scores (Table 3). We noted strongest correlations with 2-year SF-12 PCS, demonstrating that relief of pain and improved function are associated with improved physical health and physical well-being. There was negligible correlation with SF-12 MCS either at baseline or at 2-years.

We next assessed the association of pain relief and improved function with improved quality of life and compared the scores by quartiles. We compared the highest vs. the lowest quartiles for 2-year HOOS/KOOS pain and function with the change in QOL, mean SF-12 PCS and MCS scores. We found that the highest levels of pain relief and improved function were associated with the most improved quality of life, and physical well-being (Appendix Table 3). There were overall moderate correlations between 2-year HOOS/KOOS pain and function, EQ-5D, and SF-12 PCS at 2-years with QOL, but not SF-12 MCS (Table 4).

Floor and Ceiling effect

We noted a strong ceiling effect in the 2-year HOOS/KOOS pain and function scores, with no variance seen in the 2-year pain and function scores above the 50th percentile for THR

and no variance seen above the 75th percentile for TKR. Results for HOOS/KOOS pain and function are seen in Figure 1 a,b,c,d.

Clinically meaningful thresholds and sensitivity to change

The MCID thresholds for HOOS/KOOS pain and function subscales were 21 and 14 points for KOOS pain and function and 23 and 18 points for HOOS pain and function, respectively and thresholds for RID for each subscale (Table 5). The MCID and RID were anchored on the patient important domain of "a little" or "moderate" improvement in quality of life.

Responsiveness: Large ES and SRM were noted 2-years after TKR and THR for: KOOS pain, 2.3 and 1.9; KOOS function, 1.8 and 1.6; HOOS pain, 2.7 and 2.4; HOOS function, 2.4 and 2.1; and the SF-12 PCS, 1.5 and 1.2 for TKR and 1.9 and 1.6 for THR, confirming their sensitivity to change after TKR and THR (Table 6). There was little change after THA and TKA in the SF-12 MCS.

Discussion:

The OMERACT TJR Working Group has previously endorsed pain and function as core domains in the core domain set that also includes patient satisfaction, revision, adverse events, and death, to be measured in TJR clinical trials(5, 31). Our findings confirmed that with increasing relief of pain and functional improvement measured using HOOS/KOOS pain and ADL function subscales, patient reported quality of life and patient satisfaction also improved, and correlation of these constructs was demonstrated. We established patient acceptable thresholds for change in pain and function using patient anchors. While cut-points have been established for patient acceptable state after anterior cruciate ligament surgery and after TKR/THR(32, 33), the threshold for a meaningful change has not been previously established in TKR/THR populations with an intermediate follow-up. Additionally, one previous study using novel statistical technique for estimation of MCID found a small or negative change in KOOS pain and ADL after ACL repair to be consistent with patient-relevant improvements, raising questions about the sample, response shift, the anchor used or the statistical technique used(32–34). Our study further validates these HOOS/KOOS pain and function subscales as valid measures of respective core domains for use in TJR clinical trials, registries, and outcome studies. Our study also validates them as standalone sets, obviating the use of the entire 40 plus item questionnaires. The routine measurement of these core domains in all clinical trials using the same instruments will harmonize and permit pooling of data for large-scale analysis to improve the strength of future analyses (5). Our study identifies HOOS/KOOS pain and function subscales as valid measures of these domains.

We examined HOOS/KOOS pain and ADL function domains and sought to determine if the subscales were valid measures of core TJR domains of pain and function. We found that the psychometric properties of the subscales could be tested and validated using data collected prior to and 2-years after THA and TKA in a large arthroplasty registry, and confirmed that the instrument subscales of pain and ADL function correlate with the important patient centered domains of satisfaction and quality of life. Previous studies of HOOS/KOOS administration have shown that it is a feasible measure, and the entire questionnaire take

about 10–15 minutes to complete; the responder burden for each subscale is significantly less. The questionnaire is non-proprietary, and is available free of cost (12, 13). We are cognizant that evaluation of the entire TJR core domain set in TJR outcomes research and clinical trials needs to be feasible from a reporting burden perspective, not just each individual component. The remaining components of TJR core domain set that need valid measures are patient satisfaction, and adverse events; the latter is readily captured since revision and death are definitive end-points. The addition of a validated measure for patient satisfaction and capture of adverse events should not be a prohibitive patient respondent burden given that the composite can be completed within 10–15 minutes.

The large ES and SRM demonstrated in this TJR cohort demonstrate the importance of the change in HOOS/KOOS pain and ADL function subscales 2 years after THR/TKR. Prior studies have demonstrated stable ES and SRM for KOOS pain and ADL function up to one year(10), we were able to demonstrate a consistent meaningful response 2 years after TKR, which is a more meaningful time frame. Similarly, we demonstrated that the MCID and RID were achieved defined by a "little" or "moderate" change in patient reported quality of life. These MCID thresholds of 10–20 units on a 0–100 scale in our study are consistent with MCIDs described for other domains in the psychometric literature (34, 35). Moreover, these changes were demonstrated in a cohort with a high level of satisfaction, indicating that these measures are independent and the change in pain and ADL function subscales is important as a stand-alone measure.

This study has certain limitations. Non-response, though low, and non-differential by patient characteristics, is a study limitation. We chose the HOOS/KOOS as candidate measures due to their free availability and validation studies in other populations with knee/hip disease and some validation previously in TJR populations. There might be other measures that could also be alternate measurement tools for pain/function assessment after THR or TKR. Meaningful thresholds (MCID and RID) are estimates at best, and could vary by study setting, region/country and patient population. RID estimates may have been affected by the ceiling effect, which explains why their values were close to the MCID estimates. We noted the substantial ceiling effect with little variance for pain and function 2 years after THR and TKR. The two- year reporting period used may allow for further patient improvements than expected if collected at one year. Prosthetic survivorship has carried a higher threshold for acceptable reporting follow-up as long as five years before beings acceptable for some journals. The patient outcomes focus does not need that long a time period and two years is a reasonable threshold. In terms of usability for public policy, however, it is outside of the more commonly expected time period of one year. The data was all collected from patients in a high-volume tertiary care orthopedic hospital, where the patient population has a higher education level, is 90% white, and where high levels of satisfaction are reported. While this may decrease generalizability of the data, it should not impact the analysis of the measures.

Although the lower baseline scores did not threaten improvement and/or correlation with quality of life improvements, an analysis of the higher starting baseline scores was not performed. This would be important in that they might not have substantial improvements in pain and/or function. The findings in this study support the selection of a core outcome measurement set that includes the HOOS/KOOS subscales of pain and ADL function, as

measures of separate domains that correlate with important patient centered outcomes and to be used in clinical trials, registries, and outcomes reporting. The next step is to harmonize these findings with the HOOS, JR. and KOOS, JR. that have been the endorsed for use by CMS at the end of 9–12 months after surgery for their current bundled payment program(18–20). The MCID and RID for both HOOS/KOOS subscales were established using a similar Likert scale; this should allow for a cross-walk which will allow utilization of CMS administrative data bases. Using subscales of established instruments like the HOOS/KOOS should be considered for a core measurement set accepted as a minimum reporting baseline for future clinical trials, registries, and outcomes reporting.
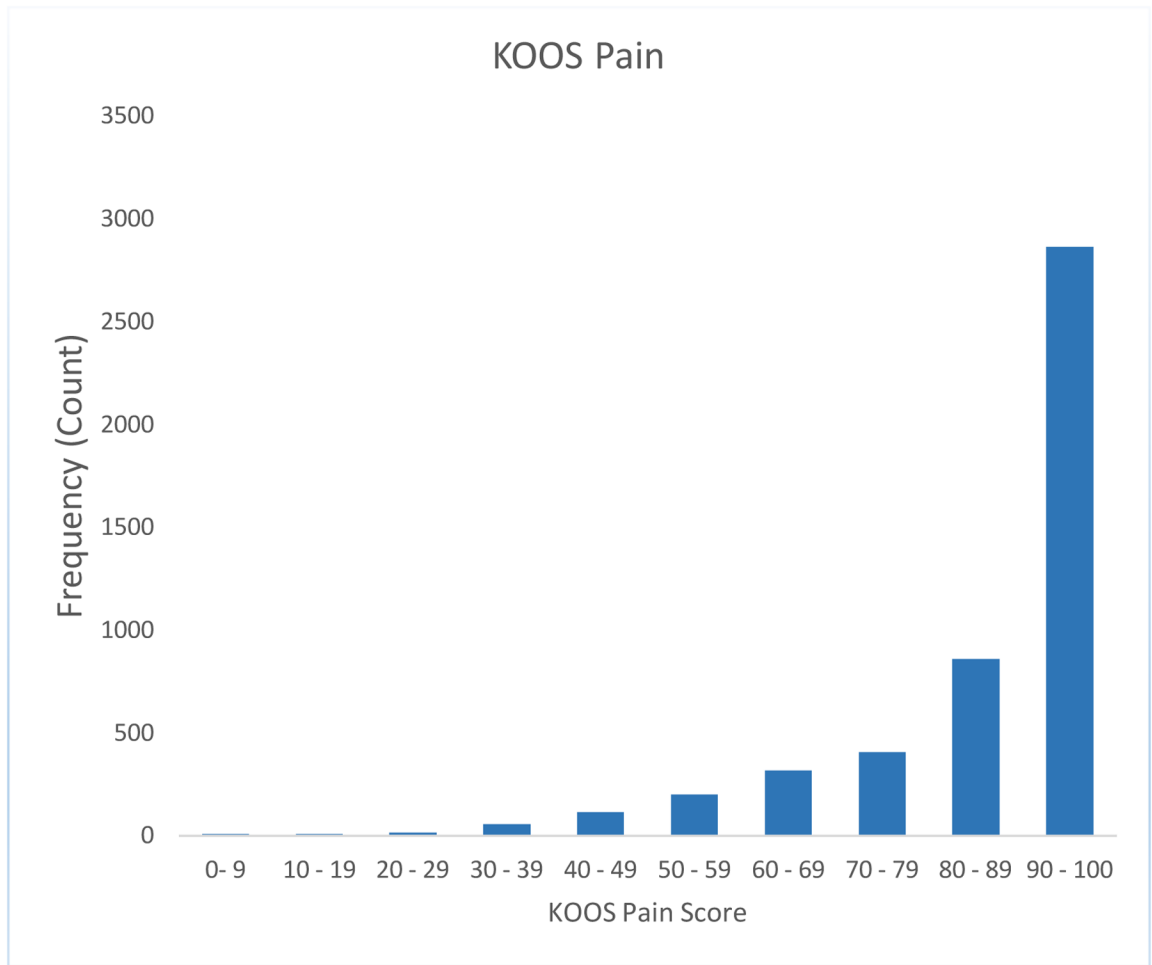
Supplementary Material

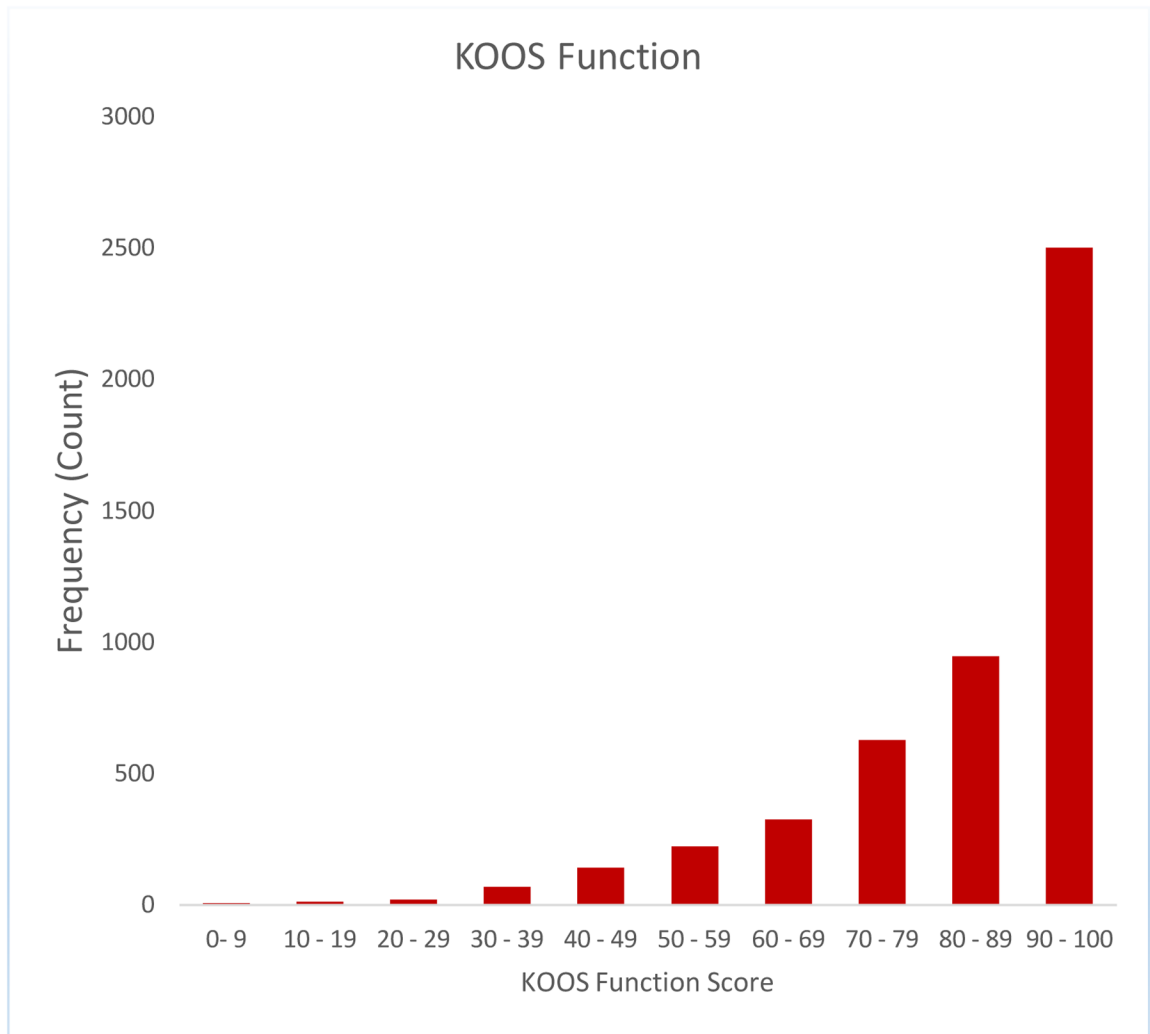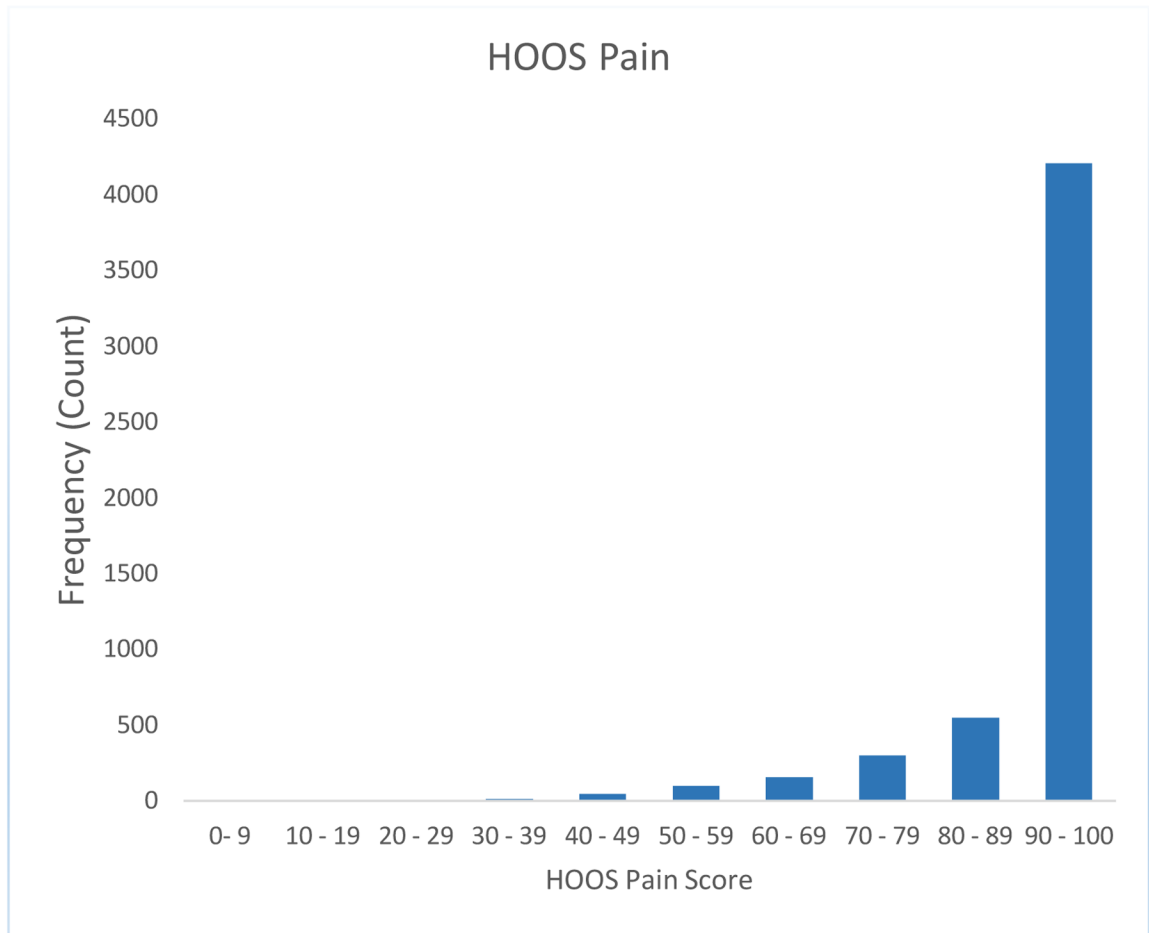Refer to Web version on PubMed Central for supplementary material.

Bibliography

1. Clement ND, Bardgett M, Weir D, Holland J, Gerrand C, Deehan DJ. Three groups of dissatisfied patients exist after total knee arthroplasty: early, persistent, and late. Bone Joint J 2018 2;100-B(2):161–9. [PubMed: 29437057]

2. Lau RL, Gandhi R, Mahomed S, Mahomed N. Patient satisfaction after total knee and hip arthroplasty. Clin Geriatr Med 2012 8;28(3):349–65. [PubMed: 22840303]

3. Riddle DL, Stratford PW, Singh JA, Strand CV. Variation in outcome measures in hip and knee arthroplasty clinical trials: a proposed approach to achieving consensus. J Rheumatol 2009 9;36(9):2050–6. [PubMed: 19738212]

4. Singh JA, Dohm M, Sprowson AP, Wall PD, Richards BL, Gossec L, et al. Outcome Domains and Measures in Total Joint Replacement Clinical Trials: Can We Harmonize Them? An OMERACT Collaborative Initiative. J Rheumatol 2015 12;42(12):2496–502. [PubMed: 25834208]

5. Singh JA, Dowsey MM, Dohm M, Goodman SM, Leong AL, Scholte Voshaar MMJH, et al. Achieving Consensus on Total Joint Replacement Trial Outcome Reporting Using the OMERACT Filter: Endorsement of the Final Core Domain Set for Total Hip and Total Knee Replacement Trials for Endstage Arthritis. J Rheumatol 2017 11;44(11):1723–6. [PubMed: 28089984]

6. Singh JA, Dowsey M, Choong PF. Patient Endorsement of the Outcome Measures in Rheumatology (OMERACT) Total Joint Replacement (TJR) clinical trial draft core domain set. BMC Musculoskelet Disord 2017 3 15;18(1):111,017–1464-x. [PubMed: 28298194]

7. Hoang A, Goodman SM, Navarro-Millan IY, Mandl LA, Figgie MP, Bostrom MP, et al. Patients and surgeons provide endorsement of core domains for total joint replacement clinical trials. Arthritis Res Ther 2017 12 6;19(1):267,017–1476-9. [PubMed: 29208013]

8. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes 2003 5 25;1:17. [PubMed: 12801417]

9. Roos EM. Joint injury causes knee osteoarthritis in young adults. Curr Opin Rheumatol 2005 3;17(2):195–200. [PubMed: 15711235]

10. Collins NJ, Prinsen CA, Christensen R, Bartels EM, Terwee CB, Roos EM. Knee Injury and Osteoarthritis Outcome Score (KOOS): systematic review and meta-analysis of measurement properties. Osteoarthritis Cartilage 2016 8;24(8):1317–29. [PubMed: 27012756]

11. Nilsdotter AK, Lohmander LS, Klassbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS)--validity and responsiveness in total hip replacement. BMC Musculoskelet Disord 2003 5 30;4:10. [PubMed: 12777182]

12. Collins NJ, Misra D, Felson DT, Crossley KM, Roos EM. Measures of knee function: I. Arthritis Care Res (Hoboken) 2011 11;63 Suppl 11:S208–28. [PubMed: 22588746]

13. Nilsdotter A, Bremander A. Measures of hip function and symptoms: Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index

of Severity for Osteoarthritis of the Hip (LISOH), and American Academy of Orthopedic Surgeons (AAOS) Hip and Knee Questionnaire. Arthritis Care Res (Hoboken) 2011 11;63 Suppl 11:S200–7. [PubMed: 22588745]

14. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol 1992 6;45(6):613–9. [PubMed: 1607900]

15. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373–83. [PubMed: 3558716]

16. Moreno RP, Pearse R, Rhodes A, European Surgical Outcomes Study (EuSOS) Group of the European Society of Intensive Care Medicine and European Society of Anaesthesiology Trials Groups. American Society of Anesthesiologists Score: still useful after 60 years? Results of the EuSOS Study. Rev Bras Ter Intensiva 2015 Apr-Jun;27(2):105–12. [PubMed: 26340149]

17. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Int J Surg 2014 12;12(12):1500–24. [PubMed: 25046751]

18. Lyman S, Lee YY, Franklin PD, Li W, Mayman DJ, Padgett DE. Validation of the HOOS, JR: A Short-form Hip Replacement Survey. Clin Orthop Relat Res 2016 6;474(6):1472–82. [PubMed: 26926772]

19. Lyman S, Lee YY, McLawhorn AS, Islam W, MacLean CH. What Are the Minimal and Substantial Improvements in the HOOS and KOOS and JR Versions After Total Joint Replacement? Clin Orthop Relat Res 2018 12;476(12):2432–41. [PubMed: 30179951]

20. Lyman S, Lee YY, Franklin PD, Li W, Cross MB, Padgett DE. Validation of the KOOS, JR: A Short-form Knee Arthroplasty Outcomes Survey. Clin Orthop Relat Res 2016 6;474(6):1461–71. [PubMed: 26926773]

21. Busija L, Pausenberger E, Haines TP, Haymes S, Buchbinder R, Osborne RH. Adult measures of general health and health-related quality of life: Medical Outcomes Study Short Form 36-Item (SF-36) and Short Form 12-Item (SF-12) Health Surveys, Nottingham Health Profile (NHP), Sickness Impact Profile (SIP), Medical Outcomes Study Short Form 6D (SF-6D), Health Utilities Index Mark 3 (HUI3), Quality of Well-Being Scale (QWB), and Assessment of Quality of Life (AQoL). Arthritis Care Res (Hoboken) 2011 11;63 Suppl 11:S383–412.

22. Brooks RG, Jendteg S, Lindgren B, Persson U, Bjork S. EuroQol: health-related quality of life measurement. Results of the Swedish questionnaire exercise. Health Policy 1991 Jun;18(1):37–48.

23. Busija L, Osborne RH, Nilsdotter A, Buchbinder R, Roos EM. Magnitude and meaningfulness of change in SF-36 scores in four types of orthopedic surgery. Health Qual Life Outcomes 2008 7 31;6:55. [PubMed: 18667085]

24. Wu M, Brazier JE, Kearns B, Relton C, Smith C, Cooper CL. Examining the impact of 11 long-standing health conditions on health-related quality of life using the EQ-5D in a general population sample. Eur J Health Econ 2015 3;16(2):141–51. [PubMed: 24408476]

25. Greene ME, Rader KA, Garellick G, Malchau H, Freiberg AA, Rolfson O. The EQ-5D-5L Improves on the EQ-5D-3L for Health-related Quality-of-life Assessment in Patients Undergoing Total Hip Arthroplasty. Clin Orthop Relat Res 2015 11;473(11):3383–90. [PubMed: 25488404]

26. Mancuso CA, Salvati EA, Johanson NA, Peterson MG, Charlson ME. Patients' expectations and satisfaction with total hip arthroplasty. J Arthroplasty 1997 Jun;12(4):387–96.

27. Mancuso CA, Sculco TP, Wickiewicz TL, Jones EC, Robbins L, Warren RF, et al. Patients' expectations of knee surgery. J Bone Joint Surg Am 2001 7;83-A(7):1005–12.

28. Vissers MM, Bussmann JB, Verhaar JA, Busschbach JJ, Bierma-Zeinstra SM, Reijman M. Psychological factors affecting the outcome of total hip and knee arthroplasty: a systematic review. Semin Arthritis Rheum 2012 2;41(4):576–88. [PubMed: 22035624]

29. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. J Exp Psychol Gen 2012 2;141(1):2–18. [PubMed: 21823805]

30. Middel B, van Sonderen E. Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. Int J Integr Care 2002;2:e15. [PubMed: 16896390]

31. Singh JA, Dowsey MM, Dohm M, Goodman SM, Leong AL, Scholte Voshaar MMJH, et al. Achieving Consensus on Total Joint Replacement Trial Outcome Reporting Using the OMERACT Filter: Endorsement of the Final Core Domain Set for Total Hip and Total Knee Replacement Trials for Endstage Arthritis. J Rheumatol 2017 11;44(11):1723–6. [PubMed: 28089984]

32. Ingelsrud LH, Terwee CB, Terluin B, Granan LP, Engebretsen L, Mills KAG, et al. Meaningful Change Scores in the Knee Injury and Osteoarthritis Outcome Score in Patients Undergoing Anterior Cruciate Ligament Reconstruction. Am J Sports Med 2018 4;46(5):1120–8. [PubMed: 29517924]

33. Connelly JW, Galea VP, Rojanasopondist P, Matuszak SJ, Ingelsrud LH, Nielsen CS, et al. Patient Acceptable Symptom State at 1 and 3 Years After Total Knee Arthroplasty: Thresholds for the Knee Injury and Osteoarthritis Outcome Score (KOOS). J Bone Joint Surg Am 2019 6 5;101(11):995–1003. [PubMed: 31169576]

34. Jones PW. Interpreting thresholds for a clinically significant change in health status in asthma and COPD. Eur Respir J 2002 3;19(3):398–404. [PubMed: 11936514]

35. 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 1;44(1):30–9. [PubMed: 24175606]

KOOS Pain

KOOS Function

Figure: Bar chart titled "KOOS Function" with x-axis "KOOS Function Score" (ranges 0-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, 90-100) and y-axis "Frequency (Count)" from 0 to 3000.
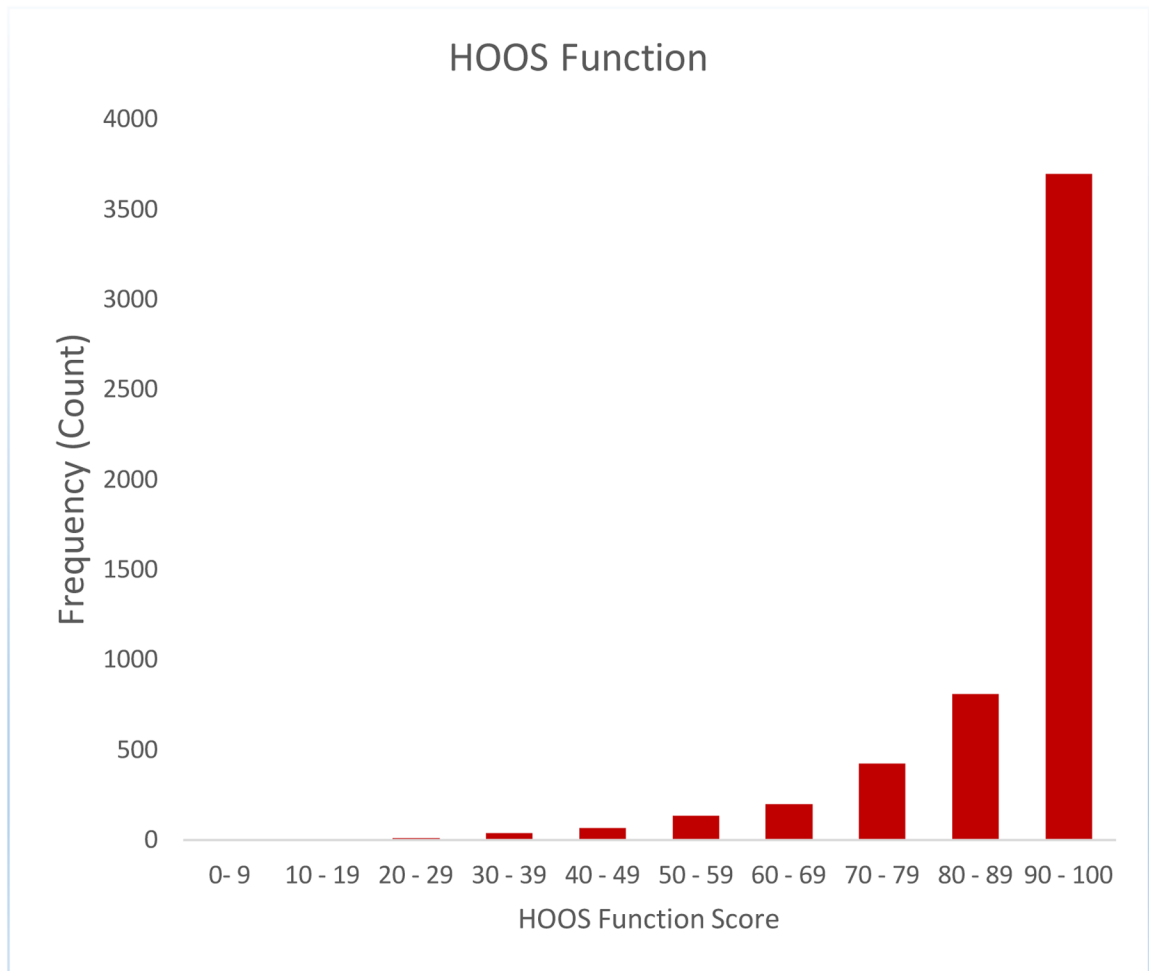
Figure 1a:
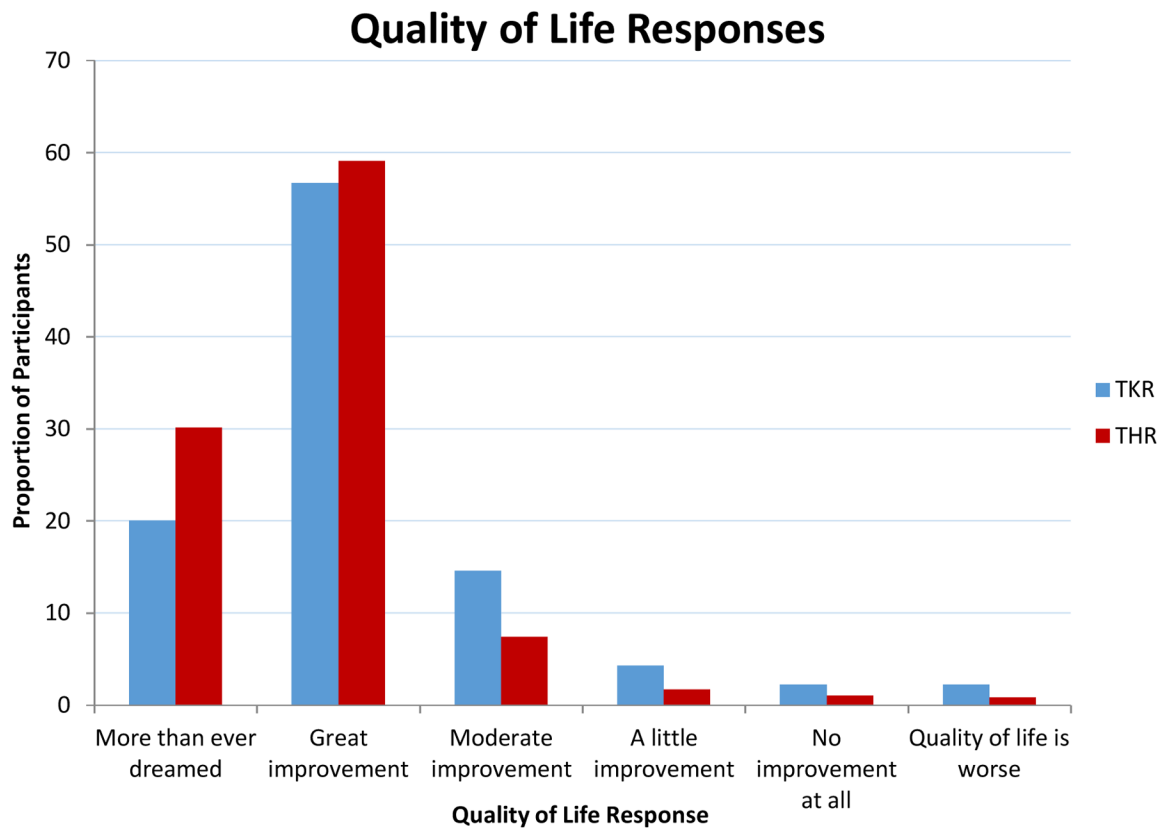A strong ceiling effect in the 2-year HOOS/KOOS pain and ADL function scores

Figure 2.
Distribution of life responses for TKR and THR pations in our cohort.

Table 1.

Characteristics of the Cohort Responder vs. non-responder characteristics (Mean, Standard Deviation)

	Overall (N=10,775)	Responders (N=10,622)	Non-Responders (N=153)	p-value
Age	**67.0 [59.0, 74.0]**	**67.0 [59.0, 74.0]**	**70.0 [63.0, 78.0]**	**<.0001**
SF-12PCS	33.3 [28.0, 38.8]	33.3 [28.0, 38.8]	33.3 [28.4, 39.2]	0.77
SF12- MCS	53.2 [42.5, 60.7]	53.2 [42.5, 60.7]	49.3 [42.0, 59.9]	0.07
Charlson Comorbidity Index	0.0 [0.0, 1.0]	0.0 [0.0, 1.0]	0.0 [0.0, 1.0]	0.02
KOOS Pain Score	47.2 [37.5, 58.3]	47.2 [37.5, 58.3]	45.8 [36.8, 62.5]	0.60
KOOS Function Score	52.9 [42.7, 66.2]	52.9 [42.7, 66.2]	50.7 [41.2, 61.0]	0.22
HOOS Pain Score	**45.0 [35.0, 57.5]**	**45.0 [35.7, 57.5]**	**39.3 [27.5, 52.5]**	**0.01**
HOOS Function Score	**48.5 [37.5, 61.8]**	**48.5 [38.2, 61.8]**	**42.7 [30.4, 61.8]**	**0.02**
Race				
Asian	53 (0.5%)	53 (0.5%)	0 (0.0%)	
Black	440 (4.1%)	430 (4.05%)	10 (6.54%)	
Hispanic	224 (2.1%)	223 (2.1%)	1 (0.65%)	0.35
Native/Pacific Islander	17 (0.2%)	17 (0.16%)	0 (0.0%)	
Other	242 (2.3%)	237 (2.23%)	5 (3.27%)	
White	9,788 (90.9%)	9,651 (90.95%)	137 (89.54%)	
Sex				
Female	6,397 (59.4%)	6,318 (59.48%)	79 (51.63%)	0.06
Male	4,378 (40.6%)	4,304 (40.52%)	74 (48.37%)	
ASA Status				
1	533 (5.0%)	530 (4.99%)	3 (1.96%)	
2	8,024 (74.5%)	7,908 (74.48%)	116 (75.82%)	0.31
3	2,204 (20.5%)	2,170 (20.44%)	34 (22.22%)	
4	10 (0.1%)	10 (0.09%)	0 (0.0%)	

*
Primary unilateral TKA and THA only. Responders includes primary unilateral THA/TKA that have any 2 year data- this includes any HOOS/ KOOS scale score or any satisfaction data at 2 years. No other restrictions applied. Bold represents statistically significant with p-value **<0.05**

SF12-PCS- Short form 12 Physical Component Score

SF12-MCS- Short Form 12-Mental Component Score

KOOS Knee injury and Osteoarthritis Outcome Score

HOOS-Hip disability and Osteoarthritis Outcome Score

ASA- Anesthesia Society of America

Table 2.

Improved Quality of Life

Response, n (%)	TKR (N = 4,641)	THR (N = 5,086)
Quality of life		
Improved More than ever dreamed	928 (20.0%)	1529 (30.1%)
Great improvement	2629 (56.7%)	3005 (59.1%)
Moderate improvement	678 (14.6%)	376 (7.4%)
A little improvement	200 (4.3%)	86 (1.7%)
No improvement at all	102 (2.2%)	51 (1.0%)
Worse	104 (2.2%)	39 (0.8%)

[*] All are 2-year HOOS/KOOS responders with QOL satisfaction data

Table 3.

Overall Spearman correlation coefficients

	TKR			THR		
	2-yr KOOS Pain Spearman ρ (95% CI)	2-yr KOOS Function Spearman ρ (95% CI)	2-yr KOOS QOL Spearman ρ (95% CI)	2-yr HOOS Pain Spearman ρ (95% CI)	2-yr HOOS Function Spearman ρ (95% CI)	2-yr HOOS QOL Spearman ρ (95% CI)
Baseline EQ-5D score	0.24 (0.21, 0.27)	0.27 (0.25, 0.30)	0.23 (0.19, 0.26)	0.20 (0.17, 0.23)	0.23 (0.20, 0.25)	0.19 (0.16, 0.23)
2-year EQ-5D score	0.53 (0.51, 0.56)	0.59 (0.56, 0.61)	0.46 (0.42, 0.49)	0.52 (0.50, 0.55)	0.57 (0.54, 0.59)	0.50 (0.47, 0.52)
Baseline MCS	0.19 (0.17, 0.22)	0.24 (0.21, 0.27)	0.21 (0.17, 0.25)	0.22 (0.19, 0.25)	0.25 (0.22, 0.28)	0.21 (0.17, 0.24)
2-year MCS	0.25 (0.22, 0.28)	0.31 (0.29, 0.34)	0.24 (0.20, 0.27)	0.27 (0.25, 0.30)	0.29 (0.27, 0.32)	0.27 (0.24, 0.31)
Baseline PCS	0.18 (0.16, 0.21)	0.24 (0.22, 0.27)	0.19 (0.15, 0.23)	0.16 (0.13, 0.18)	0.21 (0.19, 0.24)	0.20 (0.17, 0.24)
2-year PCS	0.54 (0.52, 0.56)	0.63 (0.61, 0.65)	0.54 (0.51, 0.57)	0.52 (0.50, 0.54)	0.62 (0.60, 0.64)	0.58 (0.55, 0.60)

Spearman correlation coefficients were calculated to evaluate the correlation between 2 year HOOS/KOOS pain, function and QOL with baseline and 2-year EQ-5D,SF-12 MCS, SF-12 PCS. Correlation estimates were converted to Fisher's Z to ascertain 95% confidence intervals for the corresponding estimate.

Table 4

Overall Spearman coefficients for baseline/2-year HOOS and KOOS pain and function subscale scores versus total QOL score [*]

	Knees	Hips
	QOL Spearman ρ (95% CI)	QOL Spearman ρ (95% CI)
Baseline H/KOOS pain	−0.03 (−0.06, 0.00)	0.06 (0.03, 0.09)
2-year H/KOOS pain	−0.63 (−0.64, −0.61)	−0.52 (−0.54, −0.50)
Baseline H/KOOS function	−0.06 (−0.09, −0.03)	0.02 (−0.01, 0.05)
2–year H/KOOS function	−0.63 (−0.65, −0.61)	−0.53 (−0.55, −0.52)
Baseline EQ-5D score	−0.09 (−0.13, −0.07)	0.00 (−0.03, −0.03)
2-year EQ-5D score	−0.50 (−0.52, −0.47)	−0.41 (−0.44, −0.38)
Baseline MCS	−0.10 (−0.13, −0.07)	−0.09 (−0.12, −0.06)
2-year MCS	−0.23 (−0.26, −0.20)	−0.19 (−0.22, −0.16)
Baseline PCS	−0.08 (−0.11, −0.05)	−0.01 (−0.04, −0.02)
2-year PCS	−0.49 (−0.51, −0.47)	−0.42 (−0.44, −0.40)

[*]
Note QOL scoring is 1 – best QOL, and 6 – worst QOL.

Spearman correlation coefficients were calculated to evaluate the correlation between each HOOS/KOOS pain/function subscale and total QOL score within hips and knees. Correlation estimates were converted to Fisher's Z to ascertain 95% confidence intervals for the corresponding estimate.

Table 5:

Clinically meaningful thresholds and sensitivity to change

Knee- KOOS (2y - bl)	KOOS pain, Mean (SD)		KOOS Function, Mean (SD)	
Improvement	**MCID** "A little improvement"	**RID** "Moderate improvement"	**MCID** "A little improvement"	**RID** "Moderate improvement"
Quality of life	20.96 (18.36)	29.52 (18.29)	14.20 (17.60)	23.31 (16.34)
Worsening	**MCID** "The quality of my life is worse"	---------------	**MCID** "The quality of my life is worse"	---------------
Quality of life	4.36 (22.67)	---------------	−2.86 (20.42)	---------------
Hip- HOOS (2y - bl)	**HOOS pain, Mean (SD)**		**HOOS Function, Mean (SD)**	
Improvement	**MCID** "A little improvement"	**RID** "Moderate improvement"	**MCID** "A little improvement"	**RID** "Moderate improvement"
Quality of life	22.50 (17.36)	33.85 (18.92)	18.05 (17.48)	28.54 (18.76)
Worsening	**MCID** "The quality of my life is worse"	---------------	**MCID** "The quality of my life is worse"	---------------
Quality of life	22.50 (17.36)	---------------	5.45 (28.41)	---------------

[*] Deltas for H/KOOS mean SD, according to QOL Mean Clinically Important Difference/Really Important Difference (MCID/RID) category defined by QOL response

MCID= "a little improvement"

RID= "moderate improvement"

Table 6:

Effect size and Standardized Response Mean (SRM)

	TKA effect size	TKA SRM	THA effect size	THA SRM
H/KOOS pain	2.32	1.92	2.71	2.44
H/KOOS function	1.81	1.61	2.42	2.12
PCS	1.46	1.19	1.93	1.59
MCS	0.19	0.21	0.28	0.30

*
Effect size and SRM was calculated based on the methods stated in the manuscript: effect size (ES) was calculated as mean change in scores divided by the baseline SD. Standardized response mean (SRM) was calculated as the mean change in the scores divided by the standard deviation (SD) of the changed scores.