Health Policy & Economics

Treatment Success Following Joint Arthroplasty: Defining Thresholds for the Oxford Hip and Knee Scores

Check for updates

David F. Hamilton, PhD [a, *], Fanny L. Loth, MSc [b], Deborah J. MacDonald, BA [a],
Karlmeinrad Giesinger, MD [c], James T. Patton, MD [a],
A. Hamish Simpson, DPhil, FRCS(Eng and Ed) [a], Colin R. Howie, FRCS(Ed) [a],
Johannes M. Giesinger, PhD [d]

[a] Department of Orthopaedics and Trauma, University of Edinburgh, Edinburgh, United Kingdom
[b] University of Innsbruck, Faculty of Psychology and Sport Science, Innsbruck, Austria
[c] Department of Orthopaedics, Kantonsspital St. Gallen, St. Gallen, Switzerland
[d] Innsbruck Institute of Patient-centered Outcome Research (IIPCOR), Innsbruck, Austria

ARTICLE INFO

ABSTRACT

Background: Patient-reported outcome scores are the mainstay method for quantifying success following arthroplasty. However, it is unclear when a "successful outcome" is achieved. We calculated threshold values for the Oxford Hip and Knee Score (OHS and OKS) representing achievement of a successful treatment at 12-month follow-up.
Methods: Questionnaires were administered to patients undergoing total hip (THA) or knee (TKA) arthroplasty before and 12 months after surgery alongside questions assessing key aspects of treatment success. A composite success criterion was used to perform receiver operator characteristic analysis. Thresholds providing maximum sensitivity and specificity were determined for the total sample and subgroups defined by presurgery scores.
Results: Data were available for 3203 THA and 2742 TKA patients. Applying the composite treatment success criterion, 67.3% of the TKA and 77.6% of the THA sample reported treatment success. Accuracy for predicting treatment success was high for the OHS and OKS (both areas under the curve, 0.87). For the OHS, a threshold value of 37.5 points showed highest sensitivity and specificity in the total sample, while for the OKS the optimal threshold was 32.5 points. Depending on presurgery scores, optimal thresholds varied between 32.5 and 38.5 for the OHS and 28.5 and 36.5 for the OKS.
Conclusion: This is the first study to apply a composite "success" anchor to the OHS and OKS to evaluate outcome following total joint arthroplasty. Notably fewer patients report a "successful outcome" using a composite outcome threshold than report being "satisfied."

Total joint arthroplasty is the only effective treatment to address end-stage osteoarthritis of the hip or knee. Around 100,000 hip and knee arthroplasties are carried out in the United Kingdom each year [1], while in the United States the number is >700,000 [2]. Patient-reported outcome (PRO) scores are the mainstay method for quantifying treatment success, evaluating factors such as pain, stiffness, function, satisfaction, and quality of life. Use of these metrics evaluates the success of the intervention from the patient's perspective.

Although there are various merits to evaluating outcomes with patient-reported scores, PRO measures can be challenging to interpret as they present results as score points on a range (eg, from 0 to 100 [3] or 0 to 48 [4]). These values are difficult to put into context without additional information. It is possible to use reference populations to convert these score points to standardized metrics (such as T-scores [5]) to facilitate interpretation, but while distribution-based score methods allow comparison of an individual patient or group against the reference population, this methodology does not directly evaluate the success of the intervention. Evaluation of

outcome success using PRO scores requires well-defined thresholds for treatment success that indicates at which point on the score range the result is associated with a successful outcome. Anchor-based approaches relate patient-reported scores to external criteria that denote treatment success and allow identification of thresholds (essentially cutoff values) for PRO measures that reflect these external criteria. Clearly, in an anchor-based approach, the definition and relevance of the external criterion is crucial.

Previous studies have used patient-reported satisfaction as an anchor to determine acceptable thresholds [6,7]. In recent years, satisfaction has gained interest as a single overarching metric to measure surgical outcome [8–11]; however, it is well established that patient satisfaction is influenced by factors not directly related to the surgical intervention, such as the experience of the hospital stay or socioeconomic status [12–16]; an apparently good clinical outcome does not necessarily predict a satisfied patient [17–21]. As such satisfaction may be an important but insufficient individual criterion to represent treatment success. Composite criteria incorporating multiple facets of outcome (such as pain resolution, functional ability, and satisfaction) may better represent successful outcomes. This composite criteria have previously been applied to the Western Ontario and McMaster Universities osteoarthritis index, the EuroQol- 5 dimensions, and the American Knee Society Score to explore thresholds for treatment success [22]; however, anchors and thresholds for the well-used Oxford Hip and Knee Score (OHS and OKS) have not been described.

The aim of this study was to develop thresholds that indicate postoperative treatment success 12 months after total joint arthroplasty for the OHS and OKS using a composite anchor to better reflect patient pain and functional status.

Materials and Methods

Study Design and Setting

We evaluated data from a regional joint arthroplasty registry. Before surgery, patients provided demographic indicators and completed the OHS or OKS. One year after surgery, patients answered individual questions on treatment outcomes (detailed below) and completed another OHS or OKS.

The OHS and OKS are widely used measures for the assessment of pain and functioning in patients undergoing lower limb arthroplasty [4]. The 12 questions of each joint-specific score can be answered on a 5-point response scale. The Oxford total scores range from 0 to 48, while higher scores indicate better ability.

Participants

We assessed prospectively collected data for patients undergoing lower limb joint arthroplasty at a single National Health Service teaching hospital during a 5-year period (January 2007-December 2011). The study center is the only hospital receiving adult referrals for a predominantly urban regional population of around 850,000. Data had been collected through informed consent for inclusion in a departmental database, for which regional ethical approval had been obtained (11/AL/0079).

Definition of Treatment Success

To define thresholds for treatment success measured with the OKS and OHS, we matched a previously reported composite model [22] applying the following combination of anchor items as external criterion:

(1) How satisfied are you with your operated knee/hip?

(2) How well did the surgery relieve pain in your affected joint?
(3) How well did the surgery increase your ability to perform regular activities?
(4) Would you have this operation again if it was required on another joint?

The satisfaction question could be answered on a 4-point rating scale from "very satisfied" to "dissatisfied," while the questions on pain relief and increase in functional ability could be answered on 5-point rating scales from "excellent" to "poor," plus an additional response option of "I don't know." The question on undergoing the procedure again could be answered on a 5-point rating scale from "definitively yes" to "certainly not," plus an additional response option of "I don't know."

We dichotomized the Likert scale responses to positive or negative statements (equivocal answers were considered negative). We considered the outcome to be successful if all 4 criteria were fulfilled. This strict classification system allowed the creation of the composite external criterion "treatment success."

Data from a further 2 patient questions on recommendation and expectations were available and used as sensitivity analysis; "Would you recommend this operation to someone else?" and "Have your expectations been met?" These additional questions employed the same 6-point scale response from "definitely yes" to "Certainly not."

Statistical Analysis

Sample characteristics are given as frequencies, ranges, means, and standard deviations. Spearman correlation coefficients were calculated to describe associations between the individual anchor items. To determine thresholds as predictors for treatment success, we performed a receiver operator characteristic analysis, using the aggregated variable "treatment success" as external criterion and the OHS and OKS at 12 months as well as score change since presurgery as predictors. We considered OHS and OKS values with a Youden's index (J) indicating the highest sensitivity and specificity as optimal cutoff value for the 1-year follow-up assessment and the change scores. We used the area under the curve (AUC) to determine the diagnostic accuracy of the OHS and the OKS. While an AUC value of 1.00 signifies perfectly accurate prediction of the outcome measure, a value between 0.70 and 0.80 is acceptable and values above 0.80 indicate excellent prediction [23]. For purposes of comparison, analysis was also performed separately for 3 subgroups defined by baseline OKS or OHS of each patient sample, differing in Oxford scores before surgery. We divided each sample into 3 groups (lowest, medium, highest tercile) using the 33rd percentile and the 67th percentile of the baseline score. In addition, we conducted a sensitivity analysis that included further possible external criteria to investigate the impact of definition of the composite criterion on the determination of thresholds. Statistical analysis was performed using IBM SPSS 24.0.

Results

Patient Characteristics

Preoperative and postoperative data were available for 3203 patients following total hip arthroplasty (THA) and 2742 total knee arthroplasty (TKA) patients, representing 86% of all lower limb arthroplasty procedures performed at the study center over this timeframe. Mean age of the THA sample was 68.0 (±11.3) years, and 58.4% were female. Mean OHS before surgery was 19.2 (standard deviation [SD] = 8.5) and 38.6 (SD = 9.4) 1 year after

Table 1

Sample Characteristics and Treatment Success Criteria at Presurgery and at 12-Month Follow-Up.

Variable	TKA (N = 2742)	THA (N = 3203)
Age		
Mean (SD)	70.2 (9.4)	68.0 (11.3)
Range	17.6-92.4	13.3-93.9
Sex		
Women, N (%)	1578 (57.5)	1870 (58.4)
Men, N (%)	1164 (42.5)	1333 (41.6)
Treatment success criteria		
Satisfaction	83.8%	90.1%
Pain relief	87.0%	92.3%
Increased ability	74.7%	82.8%
Surgery again	80.5%	88.8%
Total anchor	67.3%	77.6%
Lowest tercile[a]	54.9%	69.0%
Medium tercile[a]	66.7%	78.0%
Highest tercile[a]	77.8%	84.0%
Oxford score presurgery: mean (SD)	19.0 (7.5)	19.2 (8.5)
Oxford score 12 months postsurgery: mean (SD)	34.5 (10.0)	38.6 (9.4)
Recommendation to someone else	86.6%	93.7%
Expectations fulfilled	78.3%	85.6%

TKA, total knee arthroplasty; THA, total hip arthroplasty; SD, standard deviation.

[a] Grouping according to OKS/OHS baseline scores.

surgery. In the TKA sample, mean age was 70.2 (±9.4) years, and 57.5% were female. Preoperative OKS was 19.0 (SD = 7.5) and 34.5 (SD = 10.0) at 1-year follow-up. Based on the baseline scores, we categorized patients into 3 terciles as follows: OKS lowest tercile (0-14.9), OKS medium tercile (15-21.9), and OKS highest tercile (22-48); OHS lowest tercile (0-13.9), OHS medium tercile (14-22.9), and OHS highest tercile (23-48). For further details, see Table 1.

Treatment Success After TKA

In the TKA sample, 83.8% of patients reported to be satisfied or very satisfied 1 year after their joint arthroplasty (Table 1). And 87.0% reported good to excellent relief from pain and 74.7% good to excellent increased functional ability. Being asked to undergo the same procedure again, 80.5% were willing to do so. The combined (4-part) external criterion "treatment success" was met by 67.3% of the TKA patients (Table 1). Correlations between the 4 treatment success criteria were ranging from 0.55 (increased ability with surgery again) to 0.74 (pain relief with increased ability). Treatment success at 12 months was achieved in 54.9% of the patients with a preoperative OKS in the first tercile. And 66.7% of the patients in the medium tercile met the external criterion for treatment success. Finally, 77.8% of the patients in the highest tercile of baseline scores achieved our classification of treatment success.

Treatment Success After THA

Most of the patients (90.1%) reported to be satisfied with their hip arthroplasty, being relieved from pain (92.3%), and to have more functional ability (82.8%). Also, 88.8% were willing to undergo the same operation again, if needed (Table 1). In the THA sample, the external criterion "treatment success" was met by 77.6% (Table 1). Correlations between the 4 treatment success criteria were ranging from 0.49 (pain relief with surgery again) to 0.70 (pain relief with increased ability). And 69.0% of patients with a preoperative OHS in the lowest tercile reported treatment success 1 year after surgery. In the medium tercile, 78.0% of the patients met the combined criterion, while 84.0% of the patients in the highest tercile of baseline scores reported treatment success.

Thresholds for the OKS and OHS

In receiver operator characteristic analysis, we found the postoperative OKS to predict the dichotomous treatment success criterion (successful/not successful) in TKA patients with an AUC of 0.87 (Fig. 1). The cutoff "treatment success" threshold varied with preoperative OKS tercile (Table 2), with values in the 3 subgroups ranging from 0.85 (highest tercile at baseline) to 0.88 (lowest tercile at baseline).

For the OKS change score (difference between presurgery and 12-month follow-up), the AUC was 0.81 in the total sample, with values ranging from 0.81 (highest tercile at baseline) to 0.88 (lowest tercile at baseline).

In THA patients, accuracy of the prediction was 0.87 (Fig. 1), with values between 0.85 (lowest tercile) and 0.88 (medium tercile) for the subgroups defined by baseline score (Table 3). The OHS change score had an AUC of 0.78 in the total sample, with values in the respective terciles ranging from 0.82 (highest tercile at baseline) to 0.87 (medium tercile at baseline).

For the TKA sample, a cutoff value of 32.5 showed the highest sensitivity and specificity. According to the preoperative OKS, the cutoff value for the lowest tercile was 28.5. Patients in the medium tercile showed highest sensitivity and specificity at 33.5 points, while an optimal cutoff for the highest tercile was 36.5 (Table 2 and Fig. 2). For the OKS change score, optimal cutoff was a change of +17.5 points for the lowest baseline tercile, +13.5 points for the medium tercile, and +9.5 points for the highest tercile.

In the THA sample, optimal cutoff in terms of Youden's J was 37.5. Cutoffs for the subgroups were 32.5 for the lowest tercile, 37.5 for the medium tercile, and 38.5 for patients in the highest tercile at baseline (Table 3 and Fig. 3). For the OHS change score, the optimal cutoff was an improvement of +23.5 points in the lowest tercile, +15.5 points in the medium tercile, and +12.5 points in the highest tercile.

Sensitivity Analysis

To analyze the impact of the anchor composition on obtained cutoff scores, we conducted a sensitivity analysis varying the included items. We included 2 further analyses adding either the question "Would you recommend this operation to someone else?" to the composite criterion from the main analysis or the question "Have your expectations been met?" Both questions could be answered on a 6-point rating scale ranging from "definitely yes" to "certainly not" and were dichotomized for the analysis the same way as the other anchor items. In the TKA group, adding the question on fulfilled expectations resulted in the same AUC of 0.87 (95% confidence interval [CI], 0.86-0.89), while the optimal cutoff increased from 32.5 to 33.5. In the THA group, adding this question altered neither the AUC (0.87; 95% CI, 0.86-0.89) nor the optimal cutoff of 37.5 points. Adding the question on recommendation instead resulted again in an increase in the optimal cutoff point to 33.5 points in TKA patients, but did not change the AUC (0.87; 95% CI, 0.85-0.88). In THA patients, the results remained unchanged (optimal cutoff, 37.5 points; AUC = 0.87; 95% CI, 0.85-0.88). The question on recommending the procedure correlated with at least 0.48 (correlation with increased ability in THA patients) with all other anchor questions. Fulfilled expectations correlated at least with 0.54 (correlations with undergoing surgery again in THA patients).

Discussion

This study provides threshold values for identifying whether patients achieved "treatment success" following THA and TKA
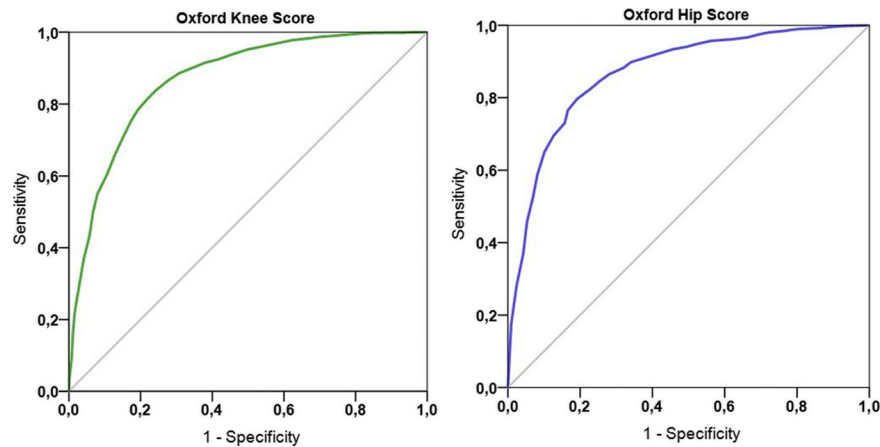
Fig. 1. ROC curves for the Oxford Knee Score (left) and Oxford Hip Score (right) predicting treatment success. ROC, receiver operator characteristic.

using the OHS and OKS; achieving 32.5 points on the OKS and 37.5 points on the OHS reflected a successful outcome.

Typically studies report that 90% of THA and 80% of TKA patients are satisfied with the outcome of surgery [13,24,25]. We required satisfaction, pain relief, functional improvement, and willingness to undergo the procedure again to be affirmed to constitute success. Applying this composite anchor resulted in approximately three-quarters (77.6%) of hip arthroplasty patients and two-thirds (67.3%) of knee arthroplasty patients reporting a successful outcome. It is currently difficult to assess the impact of new interventions and implant technologies in joint arthroplasty as, using current questionnaire assessment methodologies, most patients report high levels of outcome. As such, there is interest in deriving scoring systems without marked ceiling effects to better differentiate highly performing patients [26–28]. That this composite anchor approach "lowers" the proportion of patients reporting successful outcomes following joint arthroplasty may be beneficial as it allows for a greater "high" outcome range using existing scores.

Previous studies have proposed thresholds for the OHS and OKS by applying a single satisfaction criterion; these demonstrate notable variation in the cutoff values derived. Judge et al [7] suggested that the threshold for high satisfaction 6 months following surgery was 30 points for the OKS and 35 points for the OHS, whereas Keurentjes et al [29] reported thresholds of 37 points for the OKS and 42 points for the OHS at 3 years. This discrepancy in threshold value may be reflective of the single anchor approach,

which is susceptible to variation in response, and the studies utilized different follow-up time points.

Irrespective of the criteria employed to denote "treatment success," it may be misleading to apply the same score threshold value to all patients. We used the well-defined methodology of grouping patients by preoperative score tercile to evaluate variation in threshold value by patient case-mix factors [30]. In both hip and knee arthroplasty patients, the postoperative Oxford score required to meet the success criteria varied with preoperative tercile. A lesser absolute postoperative Oxford score was required to achieve the patient's report of "treatment success" in those with low preoperative Oxford scores.

Interestingly, the opposite pattern was observed when comparing thresholds for change in scores (between preoperation and postoperation) across the patient terciles. In patients with low baseline scores, more change was needed to qualify for treatment success than in patients with high baseline scores. The variation in threshold values derived for "change scores" was substantially larger than the variation in threshold values for absolute scores. This suggests that patients may rely on their current status as opposed to perceived improvement when reporting satisfaction postoperatively. This may reflect recall bias of their presenting symptom state. As such, postoperative absolute scores accounting for preoperative status may be more suitable for assessing treatment success than change scores.

Table 2
Accuracy of Predicting Treatment Success With the Oxford Knee Score.

	AUC	95% Confidence Interval	Cutoff	Sensitivity	Specificity
Oxford Knee Score: absolute score at 12 mo					
Total sample	0.87	0.85-0.89	32.5	0.84	0.76
Lowest tercile[a]	0.88	0.86-0.90	28.5	0.81	0.81
Medium tercile[a]	0.86	0.83-0.88	33.5	0.85	0.75
Highest tercile[a]	0.85	0.82-0.89	36.5	0.85	0.71
Oxford Knee Score: change from presurgery to 12 mo					
Total sample	0.81	0.79-0.82	+12.5	0.80	0.68
Lowest tercile[a]	0.88	0.86-0.90	+17.5	0.82	0.81
Medium tercile[a]	0.85	0.82-0.88	+13.5	0.87	0.70
Highest tercile[a]	0.81	0.77-0.85	+9.5	0.79	0.71

AUC, area under the curve.
[a] According to OKS baseline scores.

Table 3
Accuracy of Predicting Treatment Success With the Oxford Hip Score.

	AUC	95% Confidence Interval	Cutoff	Sensitivity	Specificity
Oxford Hip Score: absolute score at 12 mo					
Total sample	0.87	0.85-0.89	37.5	0.80	0.81
Lowest tercile[a]	0.85	0.83-0.88	32.5	0.80	0.77
Medium tercile[a]	0.88	0.86-0.91	37.5	0.81	0.83
Highest tercile[a]	0.86	0.82-0.89	38.5	0.90	0.69
Oxford Hip Score: change from presurgery to 12 mo					
Total sample	0.78	0.76-0.81	+16.5	0.74	0.71
Lowest tercile[a]	0.85	0.82-0.88	+23.5	0.77	0.78
Medium tercile[a]	0.87	0.85-0.90	+15.5	0.87	0.74
Highest tercile[a]	0.82	0.78-0.86	+12.5	0.71	0.79

AUC, area under the curve.
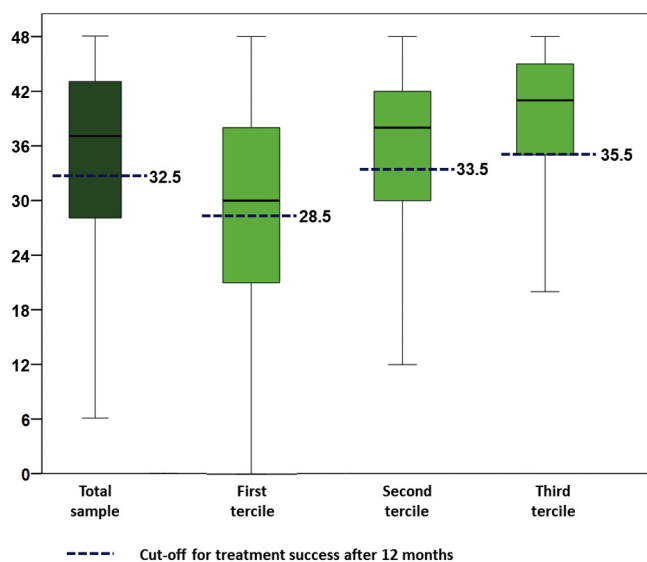[a] According to OHS baseline scores.

Fig. 2. Treatment success thresholds for the Oxford Knee Score in the total sample and in the 3 subgroups defined by presurgery scores.

Strengths and Limitations

The relatively large sample size and application of a composite criterion for treatment success can be regarded as strengths of our study. Although this is (to our knowledge) the most comprehensive anchor applied to the Oxford scores, we cannot be sure that the criteria employed encapsulates all facets of "success" following arthroplasty. Our sensitivity analysis indicated that the Oxford score thresholds we report are fairly robust toward changes to the individual anchor components. We had 6 potential anchor items available for analysis with this dataset but, however, opted to mirror the 4-item selection previously described [22] for reasons of consistency. We were able to use the additional responses to the questions of "expectations met" and "recommendation to a friend" as sensitivity analyses to explore the value of incorporating these additional factors to the composite anchor. Adding further criteria had minimal impact, with no changes to the THA threshold patients and varying the TKA threshold by 1 point. This finding reflects the
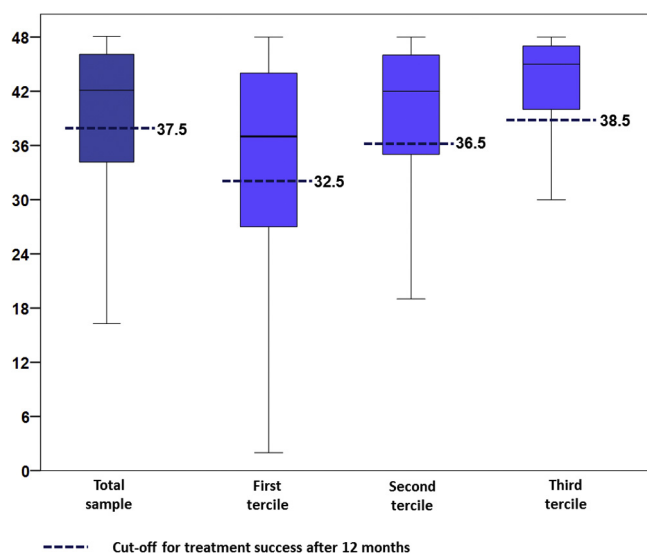

Fig. 3. Treatment success thresholds for the Oxford Hip Score in the total sample and in the 3 subgroups defined by presurgery scores.

strong correlations between the various anchor questions that may reflect overlapping content.

A limitation is that we describe thresholds for success at a single postoperative time point of 12 months. Extrapolation cannot be drawn to threshold values at different timeframes nor any longitudinal changes over the postoperative period. The 12-month postoperative time point is however the timeframe most commonly reported in arthroplasty outcome evaluation, accepted to reflect the completion of postoperative recovery.

Conclusion

This study provides thresholds for identifying "treatment success" following THA and TKA using the OHS and OKS. We employed a comprehensive definition of treatment success comprising patient satisfaction, functional improvement, pain relief, and willingness to undergo the same procedure again. Using this composite criterion, approximately three-quarters of hip arthroplasty patients and two-thirds of knee arthroplasty patients reported a successful treatment.

Acknowledgments

References

[1] National Joint Registry Editorial Board. 14th Annual Report - 2017-National Joint Registry for England, Wales, Northern Ireland and the Isle of Man. 2017. http://www.njrreports.org.uk/Portals/0/PDFdownloads/NJR%2014th%20Annual%20Report%202017.pdf [accessed 12.04.18].

[2] Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. J Bone Joint Surg Am 2007;89:780–5.

[3] Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol 1988;15:1833–40.

[4] Murray DW, Fitzpatrick R, Rogers K, Pandit H, Beard DJ, Carr AJ, et al. The use of the Oxford hip and knee scores. J Bone Joint Surg Br 2007;89:1010–4.

[5] Hamilton DF, Giesinger JM, Patton JT, MacDonald DJ, Simpson AH, Howie CR, et al. Making the Oxford hip and knee scores meaningful at the patient level through normative scoring and registry data. Bone Joint Res 2015;4:137–44.

[6] Clement ND, Macdonald D, Burnett R. Predicting patient satisfaction using the Oxford knee score: where do we draw the line? Arch Orthop Trauma Surg 2013;133:689–94.

[7] Judge A, Arden NK, Kiran A, Price A, Javaid MK, Beard D, et al. Interpretation of patient-reported outcomes for hip and knee replacement surgery: identification of thresholds associated with satisfaction with surgery. J Bone Joint Surg Br 2012;94:412–8.

[8] Hamilton DF, Lane JV, Gaston P, Patton JT, Macdonald DJ, Simpson AH, et al. Assessing treatment outcomes using a single question: the net promoter score. Bone Joint J 2014;96-B:622–8.

[9] Vissers MM, de Groot IB, Reijman M, Bussmann JB, Stam HJ, Verhaar JA. Functional capacity and actual daily activity do not contribute to patient satisfaction after total knee arthroplasty. BMC Musculoskelet Disord 2010;11:121–8.

[10] Greene ME, Rolfson O, Gordon M, Garellick G, Nemes S. Standard comorbidity measures do not predict patient-reported outcomes 1 year after total hip arthroplasty. Clin Orthop Relat Res 2015;473:3370–9.

[11] Rolfson O, Bohm E, Franklin P, Lyman S, Denissen G, Dawson J, et al. Patient-reported outcome measures in arthroplasty registries report of the patient-reported outcome measures working group of the International Society of Arthroplasty Registries Part II. Recommendations selection, administration, and analysis. Acta Orthop 2016;87 Suppl 1:9–23.

[12] Larsson BW, Larsson G, Chantereau MW, von Holstein KS. International comparisons of patients' views on quality of care. Int J Health Care Qual Assur Inc Leadersh Health Serv 2005;18:62–73.

[13] Hamilton DF, Lane JV, Gaston P, Patton JT, Macdonald D, Simpson AH, et al. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. BMJ Open 2013;3. https://doi.org/10.1136/bmjopen-2012-002525.

[14] Mira JJ, Tomas O, Virtudes-Perez M, Nebot C, Rodriguez-Marin J. Predictors of patient satisfaction in surgery. Surgery 2009;145:536–41.

[15] McLafferty RB, Williams RG, Lambert AD, Dunnington GL. Surgeon communication behaviors that lead patients to not recommend the surgeon to family members or friends: analysis and impact. Surgery 2006;140:616–24.

[16] Barrack RL, Ruh EL, Chen J, Lombardi Jr AV, Berend KR, Parvizi J, et al. Impact of socioeconomic factors on outcome of total knee arthroplasty. Clin Orthop Relat Res 2014;472:86–97.

[17] Heck DA, Robinson RL, Partridge CM, Lubitz RM, Freund DA. Patient outcomes after knee replacement. Clin Orthop Relat Res 1998:93–110.

[18] Scott CE, Bugler KE, Clement ND, MacDonald D, Howie CR, Biant LC. Patient expectations of arthroplasty of the hip and knee. J Bone Joint Surg Br 2012;94:974–81.

[19] Anderson JG, Wixson RL, Tsai D, Stulberg SD, Chang RW. Functional outcome and patient satisfaction in total knee patients over the age of 75. J Arthroplasty 1996;11:831–40.

[20] Hamilton DF, Gaston P, Simpson AH. Is patient reporting of physical function accurate following total knee replacement? J Bone Joint Surg Br 2012;94:1506–10.

[21] Terwee CB, van der Slikke RM, van Lummel RC, Benink RJ, Meijers WG, de Vet HC. Self-reported physical functioning was more influenced by pain than performance-based physical functioning in knee-osteoarthritis patients. J Clin Epidemiol 2006;59:724–31.

[22] Giesinger JM, Hamilton DF, Jost B, Behrend H, Giesinger K. WOMAC, EQ-5D and Knee Society score thresholds for treatment success after total knee arthroplasty. J Arthroplasty 2015;30:2154–8.

[23] Hosmer D, Lemeshow S. Applied logistic regression. New York: John Wiley & Sons; 1989.

[24] Baker PN, van der Meulen JH, Lewsey J, Gregg PJ. The role of pain and function in determining patient satisfaction after total knee replacement. Data from the National Joint Registry for England and Wales. J Bone Joint Surg Br 2007;89:893–900.

[25] Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KD. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? Clin Orthop Relat Res 2010;468:57–63.

[26] Hamilton DF, Giesinger JM, MacDonald DJ, Simpson AH, Howie CR, Giesinger K. Responsiveness and ceiling effects of the Forgotten Joint Score-12 following total hip arthroplasty. Bone Joint Res 2016;5:87–91.

[27] Ostendorf M, van Stel HF, Buskens E, Schrijvers AJ, Marting LN, Verbout AJ, et al. Patient-reported outcome in total hip replacement. A comparison of five instruments of health status. J Bone Joint Surg Br 2004;86:801–8.

[28] Dawson J, Beard DJ, McKibbin H, Harris K, Jenkinson C, Price AJ. Development of a patient-reported outcome measure of activity and participation (the OKS-APQ) to supplement the Oxford knee score. Bone Joint J 2014;96-B:332–8.

[29] Keurentjes JC, Van Tol FR, Fiocco M, So-Osman C, Onstenk R, Koopman-Van Gemert AW, et al. Patient acceptable symptom states after total hip or knee replacement at mid-term follow-up: thresholds of the Oxford hip and knee scores. Bone Joint Res 2014;3:7–13.

[30] Escobar A, Gonzalez M, Quintana JM, Vrotsou K, Bilbao A, Herrera-Espineira C, et al. Patient acceptable symptom state and OMERACT-OARSI set of responder criteria in joint replacement. Identification of cut-off values. Osteoarthritis Cartilage 2012;20:87–92.