WOMAC, EQ-5D and Knee Society Score Thresholds for Treatment Success After Total Knee Arthroplasty

Johannes M. Giesinger, PhD [a], David F. Hamilton, PhD [b], Bernhard Jost, MD [c], Henrik Behrend, MD [c], Karlmeinrad Giesinger, MSc, MD [c]

[a] Department of Psychiatry and Psychotherapy, Medical University of Innsbruck, Innsbruck, Austria
[b] Department of Orthopaedic Surgery, University of Edinburgh, Edinburgh, UK
[c] Department of Orthopaedics and Traumatology, Kantonsspital St. Gallen, St. Gallen, Switzerland

ARTICLE INFO

ABSTRACT

Our study aimed at developing clinical thresholds (cut-off scores) for the Western Ontario and McMaster Universities (WOMAC) osteoarthritis index, EQ-5D and Knee Society Score for discriminating between patients with and without treatment success following total knee arthroplasty (TKA). We performed a retrospective analysis of 1055 patients 2 months after TKA and 765 patients 1 year after TKA. We considered treatment successful if the patient reported high levels of satisfaction and pain relief, functional increase, and a willingness to undergo the same procedure again. Based on this criterion we identified cut-off scores that will facilitate interpretation of the WOMAC, the EQ-5D and the KSS in TKA patients.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Patient-reported outcome (PRO) measures have become a cornerstone of outcome assessment after joint surgery. A range of validated questionnaires are available for assessing joint-specific parameters (e.g. function, pain) [1–6] and general health outcomes (e.g. quality of life) [7,8]. This extension of outcome assessment reflects the fact that parameters such as any-cause-revision rate are not sufficient to provide the full picture of outcome of joint surgery. An increasing number of total knee arthroplasty (TKA) registries have therefore been supplemented with PRO measures to capture function, pain or patient satisfaction [9–12].

However, whereas an increasing number of outcome parameters are being assessed, a generic definition of treatment success after TKA is lacking. Because TKA is elective surgery, in recent years, patient satisfaction has gained interest as a single overarching outcome parameter [13–15]. However, it is well known that patient satisfaction is influenced substantially by factors not directly related to TKA, such as patient's mental status [16], hospital experience [17], cultural background [18], socioeconomic status [19] and body mass index [20]. Whereas patient satisfaction is definitely a key parameter of outcome

assessment after TKA, relying solely on this construct may not be sufficient for comprehensive treatment evaluation.

Previous attempts to define treatment success on the basis of PRO scores mostly focused on change rates, comparing pre-surgical scores with follow-up assessments, and investigated minimal important differences for commonly used PRO instruments to identify patients with and without treatment success [21–24].

However, relying only on change rates ignores the fact that patients can experience substantial improvement following surgery but still have relevant functional impairments or pain, which would suggest the inappropriateness of classifying the treatment as successful [25].

To guide interpretation of absolute scores, two methodological approaches are generally applicable: A distribution-based approach relates individual PRO scores or group means to reference data, e.g., score distributions in a general population. This allows for evaluation of the extent to which a patient recovers to 'normal' levels post-surgery (e.g. with regard to function or pain). Relating scores to reference populations does not per se provide thresholds for treatment success, but certainly improves interpretability of PRO scores. In contrast, anchor-based approaches relate PRO scores to external criteria for treatment success and allow identification of thresholds (i.e. cut-off values) for PRO measures that reflect these criteria. In an anchor-based approach, the definition of the external criterion is crucial.

In our study, we employed a rather comprehensive definition of treatment success comprising patient satisfaction, functional improvement, pain relief, and willingness to undergo the same procedure again. On the basis of this conceptualization of treatment success, we

investigated the respective thresholds for two PRO measures widely used in the orthopedic field: the joint-specific Western Ontario and McMaster Universities (WOMAC) osteoarthritis index [1] and the generic EQ-5D questionnaire [7]. In addition, we determined treatment success thresholds for the Knee Society Score (KSS), a hybrid measure including patient-reported and clinician-rated outcomes [3].

Patients and Methods

Sample

Our retrospective analysis was based on data available from the local TKA registry at the Kantonsspital St. Gallen (Switzerland). Consecutive patients from February 2006 to December 2013 that underwent primary TKA and provided PRO questionnaires at 2- or 12-month follow-up were included in the study. All knee arthroplasties were LCS complete (Low Contact Stress Knee System DePuy) knee protheses (rotating platform). Patients undergoing revision surgery within 12 months after primary TKA were excluded from the analysis ($n = 9$). All patients included in our registry provide written informed consent for anonymized data analysis. Approval for registry data analysis was obtained from the local ethics committee.

Outcome Measures

Definition of Treatment Success

As an external criterion for defining thresholds for the PRO measures, we used a combination of the following anchor questions:

- How satisfied are you with your knee arthroplasty?
  (very highly or highly satisfied vs. moderately, minimally or not at all satisfied)
- If you had the choice, would you undergo the procedure again under the same conditions?
  (yes vs. no)
- Did the surgery increase your functional capacity?
  (yes vs. no)
- Did the surgery relief your pain?
  (yes vs. no)

We considered TKA successful only if the patient fulfilled all four criteria, i.e., if the patient reported pain relief, functional improvement, high or very high satisfaction, and willingness to have TKA surgery again. This strict definition allowed creation of a dichotomous external criterion for receiver operator characteristic (ROC) analysis to determine treatment success thresholds for the WOMAC, the KSS and the EQ-5D.

WOMAC

The WOMAC Osteoarthritis Index developed by Bellamy et al [1] is one of the most commonly used, patient-reported outcome measures in patients with lower limb osteoarthritis. The questionnaire contains 24 items covering three dimensions: pain (5 items), stiffness (2 items), and function (17 items). The WOMAC has been extensively tested for validity, reliability, feasibility, and responsiveness to change over time [1,26–28]. The WOMAC scores can be linearly transformed to a 0–100 scale, with higher scores indicating more severe impairment.

EQ-5D

The EQ-5D-3L is a generic five-item questionnaire for the assessment of self-reported general health [7]. It is widely used in various fields of medical research to collect quality-of-life scores as a basis for determining health state utilities, which allow calculation of quality-adjusted life-years [29].

Knee Society Score (KSS)

The Knee Society Score [3] is a widely used, clinician-reported outcome score with good published validity data [30]. The clinician-rated portion (Knee Score) of the KSS covers pain, range of movement, alignment, and stability. The patient-reported portion (Function Score) of the KSS covers the patient's mobility (walking distance and stairs) and potential walking aids. Score range of the KSS is from 0 to 100 points for each portion, with higher scores indicating better outcome.

Statistical Analysis

Sample characteristics are given as means, standard deviations, ranges, and frequencies.

Determination of thresholds was based on ROC analyses using the outcome measures (WOMAC, EQ-5D, KSS) as predictors and the previously defined dichotomous variable 'treatment success' as the criterion. In the ROC analysis, the area under the curve (AUC) is a measure of diagnostic accuracy, i.e., the ability of an outcome measure to predict the criterion. An AUC of 0.50 equals chance, whereas 1.00 reflects perfectly accurate prediction. In fact, the AUC gives the probability that a positive case (patient with treatment success according to the external criterion) has a higher PRO score than does a negative case (a patient without treatment success). In line with Hosmer and Lemeshow [31], an AUC between 0.70 and 0.80 indicates acceptable discrimination and an AUC above 0.80 indicates excellent discrimination.

We report cut-off values for an outcome measure as a threshold for treatment success that provides the highest sensitivity and specificity (i.e., the cut-off with the highest sum of sensitivity and specificity). For purposes of comparison, we provide AUCs not only for the aggregated external criterion as defined above, but also for individual components of the criterion. Analysis was performed separately for 2- and 12-month follow-up to provide thresholds for treatment success for both time points and to investigate changes in cut-off values over time.

Results

Patient Characteristics

Analysis of treatment success 2 months post-surgery included 1055 cases (mean age: 68.8 years; 60.2% were female). Twelve months post-surgery, 765 patients were eligible for analysis (mean age: 68.4 years; 61.4% female). The two samples did not differ significantly with regard to age, sex, side of implant and body mass index (all $P > 0.30$). WOMAC, EQ-5D and KSS scores were significantly different (all $P < 0.001$) at 2- and 12-month follow-up, with the largest difference found for the KSS Function Score (effect size: Cohen's $d = 1.05$) and the smallest for the EQ-5D ($d = 0.30$). For further details see Table 1.

Treatment Success 2 and 12 Months Post-surgery

Satisfaction rates did not differ significantly between the two follow-up time points, with 77.5% of patients being very highly or highly satisfied at 2 months and 76.8% at 12 months ($P = 0.247$). In line with this, we found the same number of patients (89.6%) willing to undergo the same surgery again ($P = 0.402$) at both follow-up time points. The number of patients reporting improved function increased from 69.0% at 2 months to 83.4% at 12 months ($P < 0.001$), and the number of patients reporting less pain after surgery increased from 84.7% to 91.2% over the same period ($P < 0.001$, Table 1).

The combined (four-part) external criterion was met by 61.4% of patients at 2 months and by 70.6% at 12 months (increase statistically significant with $P = 0.001$, Table 1). Correlations between the four parts of the external criterion were between $r = 0.41$ and $r = 0.55$ at 2 months and between $r = 0.48$ and $r = 0.55$ at 12 months (all $P < 0.001$).

Thresholds for Treatment Success at 2- and 12-Month Follow-Up

The highest accuracy for predicting treatment success at 2 months was found for the WOMAC Pain score (AUC = 0.76), the WOMAC