The Munich Knee Questionnaire: Development and Validation of a New Patient-Reported Outcome Measurement Tool for Knee Disorders

Marc Beirer, M.D., Nico Fiedler, C.M., Stephan Huber, M.D., Marcus Schmitt-Sody, M.D., Stephan Lorenz, M.D., Peter Biberthaler, M.D., and Chlodwig Kirchhoff, M.D.

Purpose: To develop and validate an all-purpose patient-reported outcome questionnaire for a patient-based follow-up examination regarding knee disorders. **Methods:** Each scale of the Knee Injury and Osteoarthritis Outcome Score (KOOS), International Knee Documentation Committee (IKDC) score, Lysholm knee score, Western Ontario Meniscal Evaluation Tool (WOMET) score, and Tegner score was analyzed, and after matching of the general topics, the dedicated items underwent a fusion to the final Munich Knee Questionnaire (MKQ) item and a score comprising 33 items was created. In a prospective clinical study, we evaluated validity, reliability, and responsiveness in 152 physical active patients (75 women and 77 men; mean age, 47 years) with traumatic as well as degenerative knee disorders. **Results:** Test-retest reliability was substantial, with intraclass correlation coefficients of at least 0.91. Construct validity and responsiveness were confirmed by correlation coefficients of 0.78 to 0.86 ($P = .01$) and 0.41 to 0.71, respectively. Correlation coefficients of the original scores (KOOS, IKDC, Lysholm, WOMET, and Tegner) and the scores calculated from the MKQ were between 0.80 and 0.91 ($P = .01$). **Conclusions:** The MKQ is a reliable and valid patient-reported outcome questionnaire for assessing knee function. It seems to enable the calculation of the original items of the KOOS, IKDC score, Lysholm knee score, WOMET score, and Tegner score. **Clinical Relevance:** The MKQ facilitates the comparison of treatment results in knee disorders and allows the evaluation of treatment efficacy. Identified inadequate treatment concepts could be eliminated, leading to increased patient satisfaction and optimized quality of health care.

Patient-reported outcome (PRO) questionnaires in general are suitable for outcomes research in patients with knee disorders because of their high validity as additional tools to clinician-assessed parameters for a comprehensive evaluation of clinical outcome.[1-3] No need to travel long distances for follow-up examinations in outpatient clinics and no need for the presence of a physician reduce the logistic effort and increase the response rate, leading to an increasing use of self-assessment questionnaires in outcomes research.[4]

In this context numerous authors have reported on outcome measurement tools after knee surgery.[5-11]

However, there is no consensus in the literature about a universal tool for evaluation of knee function.

Because of the large number of questionnaires focusing on knee disorders, the direct comparability of different follow-up studies is limited. Consequently, the risk of maintaining inadequate treatment concepts is increased, leading to reduced treatment quality and decreased patient satisfaction.[12]

The purpose of this prospective study was to develop and validate an all-purpose PRO questionnaire for a patient-based follow-up examination regarding knee disorders. We hypothesized that the Munich Knee Questionnaire (MKQ) would allow for a qualitative self-assessment of the Knee Injury and Osteoarthritis Outcome Score (KOOS),[6] International Knee Documentation Committee (IKDC) score,[13] Lysholm knee score,[14] Western Ontario Meniscal Evaluation Tool (WOMET) score,[11] and Tegner score.[15]

Methods

Development of Questionnaire

A systematic review of the literature was performed to identify valid and commonly used instruments

regarding follow-up examinations in the field of knee disorders. The PubMed Web site was searched for knee-specific terms (knee, surgery, joint, lower extremity) combined with psychometric (follow-up, validity, reliability, responsiveness) and instrument-specific terms (self-reported, patient-based, measurement tool, outcome measure, questionnaire). The KOOS, IKDC score, Lysholm knee score, WOMET score, and Tegner score were identified as frequently used and valid measurement tools. To achieve a comprehensive tool, evaluation of the knee function activity of the patient seems to be crucial. Therefore the Tegner score, as a frequently used instrument to assess the activity level, was included.

To ensure content validity of the MKQ, each scale (KOOS, IKDC score, Lysholm knee score, WOMET score, and Tegner score) was analyzed for items addressing either general topics or specific items. Subsequently, matching of the general topics was performed, and the dedicated items underwent a fusion to the final MKQ item. Specific items that had no corresponding item in the other scales (e.g., instability) were incorporated without further modification. Finally, the items were allocated to 5 domains, and the weighting of each domain—in terms of the number of items allocated to each domain—was done in parallel to the previous scales.

The MKQ contains 33 items addressing 5 domains: 7 items addressing physical symptoms, 6 items addressing pain, 5 items addressing activities of daily life (ADL), 6 items regarding sports and recreational activities (SRA), and 9 items addressing physical function and knee-related quality of life (QOL) (Appendix 1, available at www.arthroscopyjournal.org). The best and least symptomatic score for each item is 0; the worst is 10. The overall score is converted to a scale of 100%, where a value of 100% indicates an excellent result and a value of 0% indicates a poor result. Typical functional abilities (flexion/extension) are depicted as photographs.

Patient Collective

All patients who presented during the consultation hours of the senior author (C.K.) at our institution between August 2012 and July 2014 because of traumatic ligament, meniscal, or bony disorders of the knee joint, as well as degenerative knee joint disorders, were asked to participate in our study. The exclusion criteria were patients with limited legal capacity or under legal supervision and patients with psychiatric diseases, dementia, or other cognitive diseases. The study protocol was approved by our local ethics committee.

Testing and Evaluation of Measurement Qualities

Floor and Ceiling Effects. According to McHorney and Tarlov,[16] floor and ceiling effects exist if more than 15% of patients achieve the highest or lowest possible score. Therefore we would define floor or ceiling effects to be present if more than 15% of our patient collective achieved the highest (100 points) or lowest (0 points) possible score of the MKQ.

Internal Consistency. Internal consistency is defined by the degree of interrelation among the tested items.[17] The subscales are based on a reflective model in which all items are defined by a manifestation of the same underlying construct. Similar to previous studies,[18] our study calculated the Cronbach α per subscale and a score above 0.70 was considered to indicate sufficient homogeneity of the subscale's items.[18,19]

Test-Retest Reliability. Test-retest reliability is defined as the extent to which scores of the same patients under the same conditions coincide on repeated measurements.[17] The period between repeated measurements should be long enough to prevent recall of the tested items but short enough to ensure that no change in the clinical symptoms has occurred.[19] In this study a period of 1 week after the initial examination was chosen to assess test-retest reliability. Intraclass correlation coefficients (ICCs) were calculated, and positive reliability was assumed when the ICC was at least 0.70 for all tested subscales.[19]

Construct Validity. Construct validity is defined as the degree to which the scores of a PRO instrument are consistent with the a priori hypothesis, based on the assumption that the PRO instrument validly measures the construct to be measured.[17] Construct validity was assessed by correlating the subscales of the MKQ with the subscales of the KOOS. The KOOS, as an extension of the Western Ontario and McMaster Universities Osteoarthritis Index, was used because it presents a valid, reliable, and responsive self-administered instrument that can be used for follow-up examinations of several types of knee injury.[20,21] Pearson correlation coefficients (PCCs) were calculated. Similar to previous studies, positive construct validity was assumed when the PCC was at least 0.70 for all measured subscales.[22]

Responsiveness. Responsiveness is defined as the ability of a PRO instrument to detect changes over time of the construct to be measured.[17] Responsiveness was evaluated 4 months after the initial presentation of patients in our outpatient clinic. To assess responsiveness, patients completed the MKQ and a global perceived effect (GPE) score consisting of only 1 question on the patients' subjective opinion regarding improvement or worsening of their knee function during the past 4 months. The list of potential answers contained 7 categories (much better [+3], better [+2], somewhat better [+1], no change