The Journal of Arthroplasty 28 (2013) 418–422

Contents lists available at SciVerse ScienceDirect

The Journal of Arthroplasty

journal homepage: www.arthroplastyjournal.org

Development Of A Computer-Adaptive Version Of The Forgotten Joint Score

Johannes M. Giesinger PhD [a], Markus S. Kuster PhD, MD [b], Bernhard Holzner PhD [a], Karlmeinrad Giesinger MSc, MD [c]

[a] Department of Psychiatry and Psychotherapy, Innsbruck Medical University, Anichstr.35, Innsbruck, Austria
[b] Department of Orthopaedic Surgery, Royal Perth Hospital, University of Western Australia, Wellington Street, Perth, WA, Australia
[c] Department of Orthopaedic Surgery, Kantonsspital St. Gallen, Rorschacherstrasse 95, St. Gallen, Switzerland

ARTICLE INFO

Article history:
Received 21 May 2012
Accepted 17 August 2012

Keywords:
patient-reported outcomes
forgotten joint score
electronic data capture
computer-adaptive testing
total knee arthroplasty
total hip arthroplasty

ABSTRACT

Patient-reported outcomes (PROs) are an important endpoint in orthopedics providing comprehensive information about patients' perspectives on treatment outcome. Computer-adaptive test (CAT) measures are an advanced method for assessing PROs using item sets that are tailored to the individual patient. This provides increased measurement precision and reduces the number of items. We developed a CAT version of the Forgotten Joint Score (FJS), a measure of joint awareness in everyday life. CAT development was based on FJS data from 580 patients after THA or TKA (808 assessments). The CAT version reduced the number of items by half at comparable measurement precision. In a feasibility study we administered the newly developed CAT measure on tablet PCs and found that patients actually preferred electronic questionnaires over paper–pencil questionnaires.

Background

Patient-reported outcome (PRO) measures are widely used in orthopedic outcome research as they provide important and detailed information on patients' perception of symptoms and function in everyday life [1,2]. Currently, various modes of administration of PRO measures are in use. Most frequently, PROs are assessed via paper–pencil questionnaires, but electronic PRO questionnaire administration is increasingly employed in clinical studies and daily clinical practice [3,4].

The recently developed Forgotten Joint Score (FJS, [5] assesses patients' awareness of their knee or hip joint during activities of daily living, representing a specific but very subjective PRO measure. It was developed as we considered joint awareness a very important and highly discriminative outcome parameter especially in patients with good to excellent joint function.

Traditional PRO measures, such as the FJS, the WOMAC osteoarthritis index [6] or the SF-36 quality of life questionnaire [7] use the same questions for all patients which is unfavorable for several reasons.

First, questionnaire length poses a certain burden to a patient as these measures require a considerable number of questions to cover the whole measurement range of the outcome parameter of interest. Second, patients find themselves confronted with questions that are not appropriate to their current condition. Inappropriate questions can be inconvenient to the patients interfering with their compliance. Third, these inappropriate questions provide no or little additional information neither to the clinician nor the researcher (e.g. if a patient reports barely being able to walk, further questions on various sports activities provide little or no further information).

To overcome this limitation of traditional PRO measures, a major focus of current research activities in PRO methodology is the development of computer-adaptive test (CAT) measures [8-10]. Computer-adaptive testing (CAT) is a sophisticated method for the assessment of PROs using individually tailored sets of questions to increase measurement precision and reduce the number of questions administered to each patient. CAT requires an item bank (i.e. a set of questions and their psychometric characteristics) and an algorithm for tailoring individual sets of questions.

Based on the response to the initial item the CAT algorithm calculates a first estimate of the PRO score and selects the next most appropriate item to be administered to the patient. This procedure continues until a predefined measurement precision has been reached or a maximum number of items have been asked. Fig. 1 provides details on the various steps of the CAT procedure.

A crucial prerequisite for CAT is the availability of strictly unidimensional item sets that allow to fit complex item response theory (IRT) measurement models [11]. Since internal consistency (i.e. unidimensionality) of orthopedic questionnaires is generally very high [12] they lend themselves very well to IRT model application and CAT measure development. Consequently, a few studies have applied IRT modeling approaches to orthopedic outcome measures for

The Conflict of Interest statement associated with this article can be found at http://dx.doi.org/10.1016/j.arth.2012.08.026.

Reprint requests: Karlmeinrad Giesinger, MSc, MD, Department of Orthopaedic Surgery, Kantonsspital St. Gallen, Rorschacherstrasse 95, CH-9000 St. Gallen, Switzerland.

0883-5403/2803-0007
http://dx.doi.org/10.1016/j.arth.2012.08.026

functional status, pain, and rehabilitation outcome [13-16] including studies on various aspects of CAT measures [13,17,18].

Patients and Methods

Objectives

Our study aimed at shortening the recently published Forgotten Joint Score [5]. Therefore, we developed a computer-adaptive version from its paper–pencil form, the FJS-CAT. In detail, we addressed the following aims:

A. to develop an item response theory measurement model and an item bank for computer-adaptive testing of joint awareness
B. to derive measurement characteristics of the FJS-CAT from a large patient sample
C. to implement the FJS-CAT in a software package and evaluate its feasibility and efficiency in clinical routine (touch tablet PC)

Sample

Development of the FJS-CAT (Aim A) was based on the analysis of FJS-12 data collected at the orthopedic outpatient unit of the Kantonsspital St. Gallen (Switzerland) between August 2007 and 2011.

Inclusion criteria for patient recruitment were the following:

- primary THA or primary TKA within the last 5 years (minimum of 1 year postoperatively)
- age between 18 and 90 years
- no obvious cognitive impairments
- written informed consent

To evaluate feasibility and efficiency (Aim C) we recruited 60 patients according to the same criteria. The study was approved by the local ethics committee.

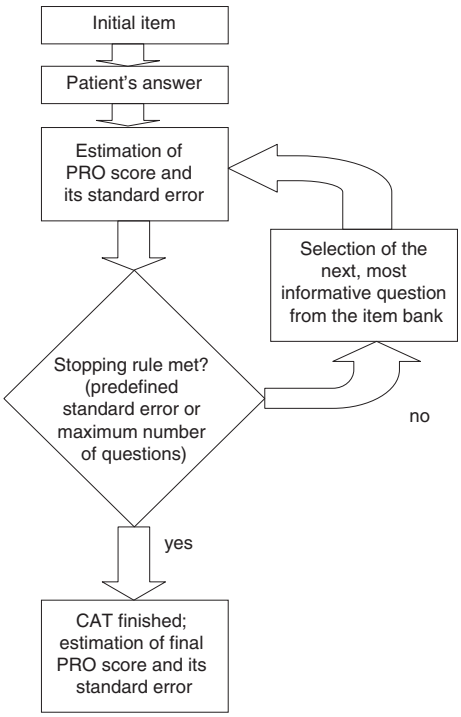

Fig. 1. Flowchart showing the CAT procedure.

The Forgotten Joint Score

Item bank development was based on the above-mentioned 12-item version of the FJS [5]. It uses a 5-point Likert response format and the raw score is transformed to range from 0 to 100 points. High scores indicate good outcome, i.e. a high degree of forgetting the joint in everyday life (forgotten joint phenomenon).

The FJS has a low ceiling effect and was designed to especially discriminate between good, very good and excellent outcome after THA and TKA. The validation study [5] showed high internal consistency (Cronbach's alpha 0.95) and the FJS score proved to discriminate well between patient groups known to show different outcome (i.e. known-group comparisons).

Psychometric and Statistical Analysis

The 12 items of the FJS were considered for inclusion in the FJS-CAT item bank and analyzed with regard to unidimensionality and fit to an IRT model. For investigation of unidimensionality we calculated Cronbach's alpha and conducted a principal component factor analysis. Analysis of fit to an IRT model was based on infit and outfit mean square statistics and root mean squared error of approximation (RMSEA). According to Linacre and Wright [19] we considered infit and outfit mean square values for individual items between 0.5 and 1.5 as an indicator of sufficient model data fit. Analysis of differential item functioning (DIF, i.e. differences in item difficulty between patient subgroups) was done with regard to sex, THA/TKA, and age (above/below 70 years). Uniform DIF was considered to be substantial if log odds ratios exceeded 0.64 with $P<.001$ and non-uniform DIF if increase in R^2 was larger than 0.035 [20].

As a measure of local independence of items we calculated inter-item residual correlations after fitting a unidimensional IRT model. According to Fliege et al. [21] residual correlations below 0.25 were considered an indicator of local independence.

We performed statistical analyses with SPSS 20.0 (SPSS Inc.), IRT analyses with Winsteps 3.42 [22] and CAT simulation with Firestar 1.2 [23].

Pilot Study On The Use Of FJS-CAT

To evaluate the feasibility of data collection with the newly developed electronic FJS-CAT we collected patient feedback within a pilot sample of 60 patients after THA or TKA. We administered the FJS-CAT version together with additional questions on usability, preferences for questionnaire administration mode and computer literacy. Time for CAT completion was recorded by the software. A software package (Computer-based Health Evaluation System (CHES) [24] was used to administer the FJS-CAT to the patients on a touch tablet PC.

Results

Patient Characteristics

FJS data from 580 patients who had undergone THA ($n=154$, 26.6%) or TKA ($n=426$, 73.4%) surgery at the Kantonsspital St. Gallen (Switzerland) were available for analysis. Mean age at surgery was 68.1 years (SD 10.4) and 56.7% were women. Data from a total of 808 assessments were available for IRT analysis. On the assessment date patients were 18.7 months after surgery on average (SD 14.4). Further details on patient characteristics are given in Table 1.

For feasibility testing of the FJS-CAT (Aim C) we recruited a sample of 60 patients (30 THA patients, 30 TKA patients). Mean age was 67.9 years (SD 10.7) and mean time since surgery was 4.2 years (SD 3.5).