Validation of a WOMAC short form for patients with hip osteoarthritis

Amaia Bilbao, José M. Quintana, Antonio Escobar, Carlota Las-Hayas, Miren Orive, Irantzu Barrio, Susana García

Introduction:
Studies in the rheumatologic or orthopaedic field use quite often information retrieved directly from the patients as main outcome. That information has been standardized in questionnaires, what are called Patient Reported Outcomes (PRO). In the case of patients with osteoarthritis of the lower limb, the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) [1] is one of the used specific health related quality of life (HRQoL) questionnaires. But, as with more PROs’ tools, length of the questionnaire conditioned the proper response of patients. It seems that, in general terms, the shorter, reliable and valid the questionnaire, the better. The aims of this study were to propose a WOMAC short form based on previously shortened versions [2,3] and to evaluate the validity and reliability of this new shortened questionnaire for patients with hip osteoarthritis undergoing total hip replacement (THR), combining classical and modern statistical techniques, such as Rasch analysis.

Methodology:
Study population: Prospective observational study of two cohorts recruited independently from various public teaching hospitals. Consecutive patients who underwent THR were eligible for the study.

Measurements: All patients on the waiting list for a THR completed the WOMAC and SF-36 questionnaires, and additional questions regarding the level of pain and function, which we will refer to as the categorical scales (classified patients as having minor, moderate, and severe pain or function), before THR. We derived the WOMAC short form (WOMAC-SF) from the original WOMAC version to evaluate pain and function in patients with hip OA, based on previously shortened versions. The pain short form included 3 items of the long form [3], and the function short form included 8 items of the long form [2]. Therefore, the WOMAC-SF that we proposed has 11 items grouped into two dimensions: pain and function. The final scores for the long and short WOMAC versions had a range of values from 0 to 100, with higher scores indicating worse health status.

Statistical Analysis: Reliability was assessed using Cronbach’s alpha coefficient [4]. A coefficient over 0.70 was considered acceptable. The construct validity was studied by means of confirmatory factor analysis (CFA) to investigate the hypothesis that the 11 items on the questionnaire addressed two factors. Different fit indexes were evaluated [5,6]: the root mean square error of approximation (RMSEA), for which a value below 0.08 was considered acceptable; the non-normed fit index (NNFI) and comparative fit index (CFI), both of which had to exceed 0.90 to be satisfactory. Factor loadings of 0.40 or higher were considered acceptable. The CFA were performed in both cohorts to study the stability of the subjacent structure of the questionnaire. Otherwise, the Rasch method was applied to the WOMAC pain and function short forms separately to ensure that the scales were unidimensional [7]. Unidimensionality was assessed by means of infit and outfit statistics, with values between 0.7 and 1.3 indicating a good fit, and through a principal components analysis (PCA) of the residuals. Unidimensionality was considered violated if, in addition to the first factors, other factors had eigenvalues exceeding 3. We evaluated the ability of the WOMAC-SF to define a
distinct hierarchy of items along the measured variable by means of an item separation index. A value of 2.0 or greater for this statistic is comparable to reliability of 0.80 and is acceptable. To detect the presence of differential item functioning (DIF), the different levels of the trait were compared by gender. Rasch analyses were performed in both cohorts to study the stability of the item logits and order across the different samples.

Convergent and discriminant validity was assessed by analysing the relationship between the WOMAC-SF domains and the SF-36 domains with the Spearman correlation coefficient. The correlation between the pain WOMAC-SF and the bodily pain SF-36 domain and between the function WOMAC-SF and the physical function SF-36 domain would be higher than with the other domains. Known-groups validation was examined by comparing the pain and function short scales among the different groups according to pain and function categorical scales. Analysis of variance using the Scheffe test for multiple comparisons was performed.

All statistical analyses were performed with SAS for Windows statistical software, version 9.1 (SAS Institute, Inc., Cary, NC), except the Rasch analysis for which we used Winsteps version 3.69.1.4 software (John M. Linacre, Chicago).

Results:
We included 788 and 445 patients in the first and second cohorts respectively. The results of the CFA for the hypothesized model of two latent factors, pain and function, provided satisfactory fit indices in both cohorts. The RMSEA values were less than 0.08, and CFI and NNFI values were all exceeding 0.90. All factor loadings were significant (P<0.001) (range, 0.53-0.84) and similar in both cohorts. Regarding the results of the Rasch analyses, items were equally ranked based on their level of difficulty in both cohorts, which supported the stability of items across the different samples. Unidimensionality was supported with infit and outfit statistics ranging between 0.7 and 1.3, with the exception of two items (1.33 and 1.32, respectively). Furthermore, the PCA of the residuals did not yield additional factors with eigenvalues exceeding 3. The person and item separation indexes exceeded 2, indicating reliability over 0.80. The presence of DIF by gender was not detected. Cronbach’s alpha coefficient was 0.74 for the pain short scale, and 0.88 for the function short scale.

As expected, the highest correlation coefficient of the pain and function short scales were found with the SF-36 bodily pain and physical functioning domains respectively (-0.48 and -0.54). Significant differences were found in both pain and function short mean scales among the three severity groups according to the pain and function categorical scales, supporting the known-groups validity.

Conclusions:
This complete validation process, which used two independent and large patient samples and combined classical and contemporary methods, showed that the 11-item WOMAC short form is valid and reliable for measuring pain and function in patients with hip osteoarthritis. Because of its simplicity and ease of application, the short form is a good alternative to the original WOMAC questionnaire and it may be useful in routine practice and clinical research.

References: