Development and Validation of the Flare-OA Questionnaire for Measuring Flare in Knee and Hip Osteoarthritis

Y. Traore, J. Epstein, E. Spitz, L. March, J.-F. Maillefert, C. Rutherford, C. Ricatte, C. Alleyrat, M. Cross, L.K. King, et al.

13 April 2021

Development and validation of the Flare-OA questionnaire for measuring flare in knee and hip osteoarthritis

Authors: Traore Youssouf[1], Epstein Jonathan[1,2**], Spitz Elisabeth[3], March Lyn[4], Maillefert Jean-Francis[5], Rutherford Claudia[6,7], Ricatte Camille[3], Alleyrat Camille[1], Cross Marita[4], King Lauren K[8], Callahan Leigh F[9], Fautrel Bruno[10,11], Buttel Thomas[12,13], Hawker Gillian[8], Hunter David J[4], Guillemin Francis[1,2]

1 Inserm, CHRUNancy, Université de Lorraine, CIC 1433 Clinical epidemiology, Nancy, France
2 Université de Lorraine, APEMAC - MICS, Nancy, France
3 Université de Lorraine, APEMAC - EPSAM, Metz, France
4 University of Sydney, Institute of Bone and Joint Research, Kolling Institute, and Royal North Shore Hospital, Rheumatology Department, Australia
5 University Hospital, Department of Rheumatology, Dijon, France
6 University of Sydney, Faculty of Science, School of Psychology, Sydney, Australia
7 University of Sydney, Faculty of Medicine and Health, Sydney Nursing School, Sydney, Australia
8 University of Toronto, Department of Medicine, Canada
9 University of North Carolina, Thurston Arthritis Research Center, Chapel Hill
10 Sorbonne Université – Assistance Publique Hôpitaux de Paris, Rheumatology Dept, Pitié-Salpêtrière Hospital, Paris, France
11 Institut Pierre Louis d'Epidémiologie et Santé Publique, Inserm UMR-S 1136, Paris, France
12 University of Sydney, Institute of Bone and Joint Research, Kolling Institute, and Royal North Shore Hospital, Rheumatology Department, Australia
13 Inner West Psychology, Sydney, Australia

Address correspondence to Dr Jonathan Epstein, MD, PhD, Laboratoire APEMAC, Campus Brabois santé, Université de Lorraine, 9, avenue de la Forêt de Haye CS 50184, 54505 Vandoeuvre-lès-Nancy, France. E-mail j.epstein@chru-nancy.fr

AbreviationsAPEMAC : Adaptation, mesure et évaluation en santé

CIC : Centre d'investigation clinique

EPSAM : Adaptation, comportements de santé et prise en charge psychologique

INSERM : Institut national de la santé et de la recherche médicale

MICS: Mesures et interventions complexes en santé

Abstract

Objective. Flares in osteoarthritis (OA) of the knee and hip (KHOA) are important for patients' daily life and clinical research. Using mixed methods, we developed a self-reporting instrument measuring flare and assessed its psychometric properties.

Methods. We constructed questionnaire items from semi-structured interviews and a focus group (patients, clinicians) by using a dual-language (English and French) approach. A Delphi consensus method was used to select the most relevant items. Patients with OA from Australia, France and the United States completed the preliminary Flare-OA, HOOS, KOOS and Mini-OAKHQOL questionnaires online. We used a factor analysis and content approach to reduce items and determine structural validity. We tested the resulting questionnaire (score 0-100) for internal consistency, convergent and known-groups validity.

Results. Initially, 180 statements were generated and reduced to 33 items in five domains (response 0=not at all, to 10=absolutely) by Delphi consensus (50 patients, 116 professionals) and an expert meeting. After 398 patients (mean [SD] age 64 [8.5] years, 70.4% female, 86.7% knee OA) completed the questionnaire, it was reduced to 19 items by factor analysis and a content approach (RMSEA=0.06; CFI=0.96; TLI=0.94). The Cronbach's alpha was >0.9 for the five domains and whole questionnaire. Correlation coefficients between Flare-OA and other instrument scores were as predicted, supporting construct validity. The difference in Flare-OA score between patients with and without flare (31.8) largely exceeded 2 SEM (5.5).

Conclusion. Flare-OA is a valid and reliable patient-reported instrument for assessing the occurrence and severity of flare in patients with KHOA in clinical research.

Keywords: Osteoarthritis - Knee – Hip –Flare – Item generation – Questionnaire – Psychometrics properties – Patient-reported outcome

INTRODUCTION

Osteoarthritis (OA), a chronic and disabling musculoskeletal disease, is one of the most common joint diseases worldwide. OA predominantly affects the knee, hip and hand joints, especially in older people[1–6]. It is responsible for pain and physical disability but can also affect mental health (anxiety and depression), sleep, work capacity, interpersonal interactions, and self-esteem, thus impairing quality of life[4]. OA is a major public health problem, ranked 9th among the causes of disability in the world in 2015[6]. Its burden continues to grow in most countries and is expected to increase with the ageing of the population[7,8].

The symptoms of OA vary among individuals and over time. Flares are often characterized by fluctuations in episodic pain of varying frequency, intensity and duration[9]. Recent literature reviews indicate that a flare in knee and hip OA is more than just an exacerbation of pain[7,10] and is identified by the quality, timing of symptoms, antecedents, and consequences[7,10,11]. It tends to be episodic, with variable duration (from a few minutes to a few hours or days), characterized by a change in treatment or behaviour[7]. OA flare can lead to intermittent use of analgesics and non-steroidal anti-inflammatory drugs as well as increased consultation. Because flare is multidimensional [7,11], the goal was to develop an appropriate multidimensional measure. OA flare is likely to cause behavioural and social disturbances[12,13]. Therefore, it encompasses several domains that an instrument should be able to capture[10,14].

Several instruments exist to capture patient-reported OA-related pain and disability[15,16]. No instrument currently measures all of the relevant domains of OA flare or takes into account the fluctuation from baseline symptom state. Given the importance of OA flare to patients with knee/hip OA, there is a need to be able to measure occurrence and severity of OA flare in knee/hip OA clinical trials assessing both existing and new therapeutics. The process has to start with creating an instrument that allows for quantifying the severity/intensity of a flare, before conducting a future work with appropriate design to delineate a threshold for clinicians to make a decision that a meaningful flare has occurred.

A project was set up as a joint international effort of healthcare professionals (HCPs), scientists and patients to develop a tool to assess OA flares more comprehensively. A preliminary definition of flare in knee and hip OA was established with experts, health psychologists, and HCPs in an Outcome Measures in Rheumatology (OMERACT) working group: *a transient state of the condition, with a duration of a few days, characterized by onset or worsening of pain, swelling, stiffness, and with associated impact on sleep, activity, functioning, and mood, that can resolve spontaneously or require adjusting therapy, even if only temporarily*[14].

The main objective of this study was to develop and validate a self-reporting instrument to assess flares of knee or hip OA.

METHODS

Study design

This study, endorsed as an OMERACT-Osteoarthritis Research Society International (OARSI) initiative, was conducted in two languages, French and English, to develop a self-reporting questionnaire that was conceptually equivalent in both languages[17]. We used mixed methods in an exploratory sequential design in two phases: 1) development of a preliminary version of the questionnaire and 2) final item selection and psychometric properties assessment. A working group (WG) of health psychologists, rheumatologists, epidemiologists, and patients supervised the study.

The National Commission for Data Protection in France (CNIL DR-2015-134) and the Ethics Committee on Human Research at the University of Sydney, Australia approved ethical and regulatory aspects. All patients and HCPs gave informed consent to participate in the research (registered at ClinicalTrials.gov: NCT02892058).

Phase 1: Development of the preliminary questionnaire

The development steps have been reported elsewhere[14]. Each step was conducted in France (French) and Australia (English), and results were studied jointly by the two teams to ensure consistency and comprehensiveness.

Step 1: Item generation

Twelve OA patients, balanced in OA joint, sex and age, underwent unstructured interviews that were analysed by two senior health psychologists to identify main themes related to a flare. Accordingly, a guide created in both languages supported semi-structured individual interviews with OA patients and HCPs and one focus group with patients. Transcripts were analysed by using the Nvivo software. Content analysis allowed for identifying verbatim statements that were meaningful for patients and HCPs.

Step 2: Item grouping by domains and domain selection

The WG removed redundant or similar items, then items were grouped by topic, and labels for each group were proposed as domain names. Subsequently, we conducted two Delphi surveys with patients and HCPs to keep or reject items by using a threshold of 70% agreement[14]. People recruited, independent from people in step 1, were individuals with OA at the Royal North Shore Hospital (Australia) and from the Knee and Hip OsteoArthritis Longitudinal Assessment (KHOALA) cohort study in France and HCPs (nurses, physiotherapists, rheumatologists, general practitioners, orthopedic surgeons, scientists) from OARSI and OMERACT.

The WG then approved or rejected the item groups and domain labels and agreed on the most relevant domains[14]. It set the scale modality for all items as a numerical rating scale from 0 to 10 (i.e., 0, not at all, to 10, absolutely)[18]. The total score obtained from the mean of all item scores was transformed to a scale from 0 (no flare) to 100 (worst flare). When more than 50% of responses in a domain were missing, the whole domain was considered missing.

Phase 2: Field testing: final item selection and psychometric properties

We tested the preliminary questionnaire online with OA patients enrolled from March 2019 to February 2020. Further cultural adaptation in other languages including Spanish are under way.

1. Study population

Patients were recruited at the Royal North Shore Hospital, Sydney (Australia), from the KHOALA cohort, five rheumatology centers in France, and the OA Action Alliance (OAAA) in the United States, a network of patients with self-declared OA. Physicians at participating centres in France and Australia confirmed the diagnosis of knee or hip OA.

The inclusion criteria were age ≥ 45 years with symptomatic clinical and radiological OA of the knee or hip, confirmed by a doctor (except in the United States), regardless of the onset date of OA, and able to complete online self-reporting questionnaires in English or French.

We estimated sample size according to classical test theory, recommending a number of subjects of seven times the total number of items to test construct validity, at least 50 subjects to explore floor and ceiling effects, and 10 subjects per item for confirmatory factor analysis (CFA). By taking into account the most demanding property, given the number of candidate items at this stage, the minimum sample size was 330[19–22].

2. Data collection

The preliminary Flare-OA questionnaire (phase 1) was included in the online survey.

The dimensions of the Hip and Knee Disability and Osteoarthritis Outcome Score (HOOS and KOOS) questionnaires are identical and include the same number of items (40): pain, symptoms, function in daily living (ADL), function in sport and recreation (Sport/Rec), and quality of life (QoL). Scores are calculated on a scale from 0 (extreme problems) to 100 (no problems). The Mini-Osteoarthritis Knee and Hip Quality of Life Questionnaire (Mini-OAKHQOL) contains 20 items in five dimensions: pain, physical activity, mental health, social support, and social functioning, as well as three independent items dealing with sex life, work

life, and fear of being dependent. It uses an numerical rating scale from 0 to 10 to score items and the mean item score for dimensions[23].

We collected the characteristics of age, sex, country of residence, and affected joint. An anchor question asked about the occurrence and duration of a flare over the last 4 weeks.

3. Statistical analysis

Descriptive statistics are expressed as number (percentage) for categorical data and mean (standard deviation) (range) for quantitative variables. We assessed data quality and completeness by the percentage of computable scale scores (>50% completed items). We determined scale-to-sample targeting by investigating whether scale scores spanned the entire scale range. We assessed floor and ceiling effects by the distribution of responses observed at the lower and upper extremes of an item. They were considered present if more than 15% of respondents had the lowest (0) or highest (10) possible response, respectively[24].

We used CFA to establish the most appropriate factor structure of the questionnaire by assessing the model fit indices[22,25]. For fit indices, a value of 0.08 is considered acceptable and 0.05 excellent for the Root Mean Square Error of Approximation (RMSEA) (26); a value of 0.05 is considered excellent for the Standardized Root Mean Square Residual (SRMR)[22,26,27]. The Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) are considered acceptable at 0.90 and good at 0.95[25]. The CFA allowed for selecting items leading to the most satisfactory model. We deleted or moved (domain change) items according to four criteria[28]: 1) change indices obtained in CFA: items associated with a high change index were more likely to be changed, and items for which multiple changes (typically multiple associations with different domains or items) were more likely to be changed or removed; 2) floor (or ceiling) effect observed; 3) content analysis of the items; and 4) preferences expressed in the Delphi survey (phase 1, step 2). Item loadings > 0.4 were considered acceptable.

We determined internal consistency of each domain (with more than two items) and the overall total score by the Cronbach's alpha. A value ≥ 0.70 is deemed acceptable[29].

We assessed convergent validity with Pearson correlation, comparing the Flare-OA score with scores for the other questionnaires. We expected a high correlation (>0.70) with the Flare-OA score for the pain, symptoms, and ADL domains of the HOOS and KOOS and pain, physical activities and mental health domains of the Mini-OAKHQOL; a medium correlation (0.50-0.70) for the other domains of the HOOS and KOOS; and a low to medium correlation (<0.50) for the other domains of the Mini-OAKHQOL.

We determined the known-groups validity by comparing patients according to response to the flare anchor question, "During the last 4 weeks, have you had a flare of your osteoarthritis?", by Student *t* test. We calculated the standard error of measurement (SEM) as the estimate of the standard deviation of the score multiplied by the square root of 1 minus the Cronbach's alpha to put the difference between groups into perspective[30].

We used SAS v9.4 (Cary, NC, USA) for analysis. P<0.05 was considered statistically significant.

RESULTS

Phase 1: Development of the preliminary questionnaire

Step 1: Item generation

From semi-structured interviews with 29 patients and 16 HCPs and one focus group with 10 patients, 180 statements in French (106) and English (77) were generated.

Step 2: Item grouping by domain and domain selection

On the basis of redundancy or similarity, 50 items with equivalent meaning in both languages grouped in nine domains were retained. After two Delphi rounds involving 50 patients and 116 HCPs from 17 countries on four continents, 24 items were retained by patients and 12 by HCPs, 12 concordantly; no items were rejected by patients, and HCPs rejected 18 based on the 70% agreement threshold, therefore none concordantly. The WG subsequently endorsed 5 domains [14] of items kept by patients or HCPs, resulting in a preliminary questionnaire of 33 items in the form of statements with response options capturing intensity or frequency. The items were

distributed in five domains assessing pain (6 items), swelling (2 items), stiffness (2 items), consequences of symptoms (14 items) and psychological aspects (9 items).

Phase 2: **Final item selection and psychometric properties**

In total, 398 patients participated in the online survey (mean (SD) age 64 (8.5) years; 70.4% women, predominantly Australian [57.0%] and with knee OA [86.7%]) (Table I). More than half (64.8%) had a flare over the past 4 weeks, with mean duration of flare of 8.2 (10.1) days.

Few questionnaire items (3.8%) (up to 15 of 398 questionnaires per item) had missing answers. We observed floor effects for all items and some ceiling effects for items "*I had more difficulty walking on uneven ground*", "*I felt frustrated because I was limited in my daily activities"* and "*I needed to avoid certain movements or activities to prevent my pain"* (Table II).

For construct validity, the initial model with 33 items had poor fit indices (RMSEA=0.09; SRMR=0.05; CFI=0.85 and TLI=0.82). All items had loadings of >0.4 in their respective domains. Some items were associated with two or more domains, and the "consequences of symptoms" domain was strongly related to the other four domains. After content considerations and index changes for the CFA, 14 items were removed. For example, six items from the consequences of symptoms domain were removed because they were highly related to quality of life and considered redundant with items from other domains, such as "*My pain prevented me from concentrating (focusing) on a task*" and "*I needed to take more pain medication than usual*". When translating to Spanish during phase 2, as an intended aim, the item "*I felt my lifestyle was altered*" had interpretation or translation problems because the term "lifestyle" translates better to "partying" in Spanish. The psychological aspects domain lost three items that overlapped with content captured by items in the consequences of symptoms domain.

In its final form, the Flare-OA questionnaire included 19 items, resulting in a significant improvement in the fit indices as compared with the initial model (RMSEA=0.06; SRMR=0.04;

CFI=0.96 and TLI=0.94). Cronbach's alpha for the five domains and complete 19-item questionnaire were 0.95-0.96 and 0.96, respectively. A score was computable for 384 (96%) patients.

For convergent validity (Table III), coefficients were from -0.60 to -0.86 for the Flare-OA score correlated with scores for pain, symptoms, and ADL domains of the HOOS and KOOS and pain, physical activites and mental health domains of the Mini-OAKHQOL. For the other domains of the HOOS and KOOS, coefficients ranged from -0.52 to -0.72. For the other domains of the Mini-OAKHQOL, coefficients ranged from -0.12 to -0.64. With increasing Flare-OA score, indicating worsening, the other scores decreased accordingly.

For known-groups validity, the mean [95%CI] score was 48.7 [45.7–51.7] for patients reporting a flare in the past 4 weeks in contrast to 16.9 [13.6–20.7] for those with no flare, with a difference (31.8) largely exceeding 2 SEM (5.5) ($p<0.0001$).

DISCUSSION

The new 19-item Flare-OA self-reporting questionnaire is the first instrument to measure the occurrence and severity of flares in knee and hip OA. Several steps in its development using mixed methods and a content-driven approach via item generation, selection, and reduction, based on an explicit definition of the phenomenon being measured, provide evidence of strong content validity.We paid particular attention to the Delphi process when selecting items and privileging expert opinion over pure statistics, so that the items explore facets of the latent concept in a complementary manner[28]. The hypothesised factor structure was supported by factor analyses, and an evaluation of psychometric properties in a large field test supported its validity and reliability. The high internal consistency of each factor satisfied requirement proposed for individual monitoring[31] (ref McHorney). The Flare-OA questionnaire will be useful to help clinicians assess flare severity/intensity. Further workwill be set up to delineate a threshold for making a decision that a meaningful flare has occurred[32].

The qualitative work in this study demonstrate that flare is more complex than simply the presence or absence of pain. To adequately assess flare, a more comprehensive assessment approach was needed, capturing the full range of components that constitute a flare.

On CFA, the questionnaire showed good construct validity, with an overall structure coherent with results of the qualitative analysis. It fit with domains generated to characterize flare comprehensively, recently endorsed by OMERACT international consensus (manuscript in preparation). Despite patients and HCPs reaching consensus to keep item "*I needed to take more pain medication than usual*", as a genuine consequence of a flare, the item was removed because the double-barrelled content (pain and medication) loaded simultaneously on two domains and decreased the structure validity. The floor effect observed for all items was not surprising because more than one third of patients reported no flare.

Although the correlation of the Flare-OA score with KOOS symptoms and ADL domain scores was slightly lower than expected, overall correlations were high between similar constructs as expected, notably between pain domains and Flare-OA scores, which give clues for the convergent validity. However, these correlations were not perfect, which makes sense because the measured concept, while related, differed.

The capacity to measure and characterize a flare in OA is reflected in a questionnaire's ability to discriminate between patients reporting a flare and no flare, with a significant difference of >30 points in the overall group scores. With a SEM of 5.5, one can reasonably conclude that the Flare-OA questionnaire can distinguish and quantify different flare states, at least at the population level.

The questionnaire was developed and tested in developed countries. Studies are underway for a cultural adaptation of the Flare-OA questionnaire in Moroccan Arabic and Turkish languages.

Future work should evaluate the validity of the response scales based on item response theory and evaluate responsiveness in longitudinal studies, for which preliminary SEM data offer promising perspectives.

To better adapt the short- and long-term treatment of pain in OA, we need to be able to detect the occurrence of OA flares, integrating aspects such as anxiety, fatigue and activity limitations[4,13,33,34]. Another patient-reported outcome, the the Intermittent and Constant Pain of Osteoarthritis (ICOAP), assesses pain intensity in the onset of OA as well as pain consequences[16]. A regular assessment of pain with instruments such as the ICOAP could be used to determine flare in OA as a sudden worsening of the score; however, such use is not always possible or practical. An instrument conceived from the start to measure the aspects of a flare in OA in the past 4 weeks might be more useful but should be evaluated in a comparative study of these questionnaires for assessing flare in OA. Because such a flare do not necessarily trigger a treatment change, be it by prescription or over the counter, it was difficult to identify a criterion indicating the end of the flare episode. So we relied on patients' report that they were in a flare state, according to the definition.

A strength of the study is that both patients and HCPs were involved in both item generation and domain definitions for this patient-reported outcome. Second, the two-language approach ensured that the same intended content and meaning were captured in both languages. Overall, developing the items in two languages limits difficulties of adaptation to languages other than French and English, notably by limiting the use of idioms and colloquialism[35], keeping a high level of equivalence, although a rigourous translation and cultural adaptation process remains required[36].

Limitations include a possible selection bias because the information system required respondents to complete the electronic version of the whole questionnaire; but it allowed for few missing data. Also, to obtain a single instrument measuring flare accurately for each joint, we needed to exclude patients with both affected joints. Thus, patients with both knee and hip

OA might be misrepresented in this study; however, the items had good acceptability for both disease types, and the content of the questionnaire was designed to be neutral to affected joint.

In conclusion, the Flare-OA questionnaire is valid and reliable and suitable for estimating the occurrence and severity of flare of the knee and hip OA in patients participating in clinical research.

Note: The Flare-OA questionnaire is available upon request from the authors.

Author conbritutions

(1) **Study conception and design:** FG, DH, ES, LM, C Rutherford, C Ricatte, BT, GH, LC, JE contributed to the study conception and design. **Acquisition data:** JFM, DH, LM, LC, GH, FG, ES, C Ricatte, CA, MC contributed to the acquisition of data. **Analysis and interpretation of data:** ES, C Ricatte, C Rutherford, TB, MC, YT, JE, FG, LK, GH, contributed to the analysis and the interpretation of data.

(2) **Drafting the manuscript:** YT and FG, JE, with substantial contribution from C Rutherford, JFM, MC, GH, DH. **Critically revisions for important intellectual content:** All authors.

(3) **Final approval** of the version to be published: All authors

(4) **Working Group (WG):** Annica Barcenilla-Wong, Thomas Buttel, Sam Michel Cembalo, Marita Cross, Francis Guillemin (co-chair), Gillian Hawker (co-chair), David J. Hunter (co-chair), Lyn March (co-chair), Ricatte Camille, Claudia Rutherford, Amandine Schoumacker, Elisabeth Spitz. **Collaborators:** Melanie Beguinet, Nancy, France and Joanna Makovey, Sydney, Australia.

Conflict of interest

Dr. Callahan reports other from Movement is Life Steering Committee, other from NIH Grant Panel Review, other from Balanced Pain Management Meeting, outside the submitted work. .

Dr. Cross reports personal fees from Up To Date, other from Bone & Joint Foundation, Global Alliance for Musculoskeletal Health, outside the submitted work.

Dr David Hunter reports personal fees from Pfizer, Lilly, Merck Serono, TLCBio, Kolon Tissuegene, outside the submitted work.

Dr. March reports personal fees from Pfizer Australia Pty Ltd, outside the submitted work. All other authors have nothing to disclose.

Funding

ClinicalTrials.gov: NCT02892058

Supported by grants from Association Française de Lutte Anti Rhumatisme – France Rhumatisme, the Fondation APICIL (France) and the Kolling Foundation (Sydney, Australia). Working group members were supported indirectly by their institutions.

Acknowledgements

For their invaluable contribution to the conduct of this study, we express our sincere gratitude to people, patients, health professionals and scientists from around the world belonging to various research centers who participated in the various stages of this work (Allison Albright, Kirsten Ambrose, Sylvia J. Bedford-Guaus, Jean-Marie Berthelot, Michel De Brandt, Philippe Gaudin, Anne-Christine Rat**,** Mariana Vitaloni), scientific societies and patient associations, including the Osteoarthritis Research Society International (OARSI) and OMERACT members.

REFERENCES

1. Palazzo C, Nguyen C, Lefevre-Colau M-M, Rannou F, Poiraudeau S. Risk factors and burden of osteoarthritis. *Ann Phys Rehabil Med*. 2016;59(3):134-38.

2. Vina ER, Kwoh CK. Epidemiology of osteoarthritis: literature update. *Curr Opin Rheumatol*. 2018;30(2):160-67.

3. Majani G, Giardini A, Scotti A. Subjective impact of osteoarthritis flare-ups on patients' quality of life. *Health Qual Life Outcomes*. 2005;3:14.

4. Abdul Kadir A, Arif M, Faizal M, Ishak A, Hassan II, Mohd Noor N. Adaptation and validation of the malay version of the osteoarthritis knee and hip quality of life questionnaire among knee osteoarthritis patients. *BioMed research international*. 2018;2018:1-15.

5. Lawrence RC, Helmick CG, Arnett FC, et al. Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States. *Arthritis & Rheumatism: Official Journal of the American College of Rheumatology*. 1998;41(5):778-99.

6. Sebbag E, Felten R, Sagez F, Sibilia J, Devilliers H, Arnaud L. The world-wide burden of musculoskeletal diseases: a systematic analysis of the World Health Organization Burden of Diseases Database. *Ann Rheum Dis*. 2019;78(6):844-48.

7. Cross M, Smith E, Hoy D, et al. The global burden of hip and knee osteoarthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis*. 2014;73(7):1323-30.

8. Safiri S, Kolahi A-A, Smith E, et al. Global, regional and national burden of osteoarthritis 1990-2017: a systematic analysis of the Global Burden of Disease Study 2017. *Ann Rheum Dis*. 2020;79(6):819-28.

9. Thomas MJ, Neogi T. Flare-ups of osteoarthritis: what do they mean in the short-term and the long-term? *Osteoarthritis and Cartilage*. 2020;28(7):870-73.

10. Parry EL, Thomas MJ, Peat G. Defining acute flares in knee osteoarthritis: a systematic review. *BMJ open*. 2018;8(7):e019804.

11. Murphy SL, Lyden AK, Kratz AL, et al. Characterizing Pain Flares From the Perspective of Individuals With Symptomatic Knee Osteoarthritis. *Arthritis Care Res (Hoboken)*. 2015;67(8):1103-11.

12. Thomas MJ, Rathod-Mistry T, Harper S, et al. Acute Flares of Knee Osteoarthritis (the ACT-FLARE Study): Protocol for a Web-Based Case-Crossover Study in Community-Dwelling Adults. *JMIR Res Protoc*. 2019;8(4):e13428.

13. Cross M, Dubouis L, Mangin M, et al. Defining flare in osteoarthritis of the hip and knee: a systematic literature review—omeract virtual special interest group. *J Rheumatol*. 2017;44(12):1920-27.

14. Guillemin F, Ricatte C, Barcenilla-Wong A, et al. Developing a Preliminary Definition and Domains of Flare in Knee and Hip Osteoarthritis (OA): Consensus Building of the Flare-in-OA OMERACT Group. *J Rheumatol*. 2019;46(9):1188-91.

15. 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(12):1833-40.

16. Hawker GA, Davis AM, French MR, et al. Development and preliminary psychometric testing of a new OA pain measure--an OARSI/OMERACT initiative. *Osteoarthr Cartil*. 2008;16(4):409-14.

17. Herdman M, Fox-Rushby J, Badia X. A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Qual Life Res*. 1998;7(4):323-35.

18. Lord FM. *Applications of Item Response Theory to Practical Testing Problems*. Routledge Ed; 1980. 273.

19. Souza AC de, Alexandre NMC, Guirardello E de B. Psychometric properties in instruments evaluation of reliability and validity. *Epidemiologia e Serviços de Saúde*. 2017;26:649-59.

20. Mokkink LB, Terwee CB, Patrick DL, et al. COSMIN checklist manual. *Amsterdam: University Medical Center*. Published online 2012.

21. Petrillo J, Cano S, McLeod L, Coon C. Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: a comparison of work... -. *Value Health*. 2015;18(1):25-34.

22. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*. 1999;6(1):1-55.

23. Guillemin F, Rat A-C, Goetz C, Spitz E, Pouchot J, Coste J. The Mini-OAKHQOL for knee and hip osteoarthritis quality of life was obtained following recent shortening guidelines. *J Clin Epidemiol*. 2016;69:70-8.

24. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.

25. Taasoobshirazi G, Wang S. The performance of the SRMR, RMSEA, CFI, and TLI: An examination of sample size, path size, and degrees of freedom. *J Applied Quantitative Methods*. 2016;11(3):31-9.

26. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. *Psycho bul*. 1980;88(3):588.

27. Browne MW, Cudeck R. Alternative ways of assessing model fit. *Sociological methods & research*. 1992;21(2):230-58.

28. Goetz C, Coste J, Lemetayer F, et al. Item reduction based on rigorous methodological guidelines is necessary to maintain validity when shortening composite measurement scales. *J Clin Epidemiol*. 2013;66(7):710-18.

29. Nunnally J. Psychometric Theory–2nd Edition–McGraw-Hill. *New York*. Published online 1978. 701.

30. Tighe J, McManus I, Dewhurst NG, Chis L, Mucklow J. The standard error of measurement is a more appropriate measure of quality for postgraduate medical assessments than is reliability: an analysis of MRCP (UK) examinations. *BMC medical education*. 2010;10(1):40.

31. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res*. 1995;4(4):293-307.

32. Myasoedova E, De Thurah A, Erpelding M-L, et al. Definition and construct validation of clinically relevant cutoffs on the Flare Assessment in Rheumatoid Arthritis (FLARE-RA) questionnaire. In: *Seminars in Arthritis and Rheumatism*. Vol 50. Elsevier; 2020:261-65.

33. Runhaar J, Zhang Y. Can we prevent OA? Epidemiology and public health insights and implications. *Rheumatology (Oxford)*. 2018;57(suppl_4):iv3-9.

34. Quintana JM, Arostegui I, Escobar A, et al. Validation of a screening questionnaire for hip and knee osteoarthritis in old people. *BMC Musculoskelet Disord*. 2007;8:84.

35. Erkut S, Alarcón O, Coll CG, Tropp LR, García HAV. The dual-focus approach to creating bilingual measures. *J Cross-Cultural Psychol*. 1999;30(2):206-18.

36. Epstein J, Santo RM, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *J Clin Epidemiol*. 2015;68(4):435-41.

Tables

Table I Sociodemographic and clinical characteristics of patients with osteoarthritis (OA) completing the Flare-OA questionnaire (n=398)

Sex	
Male	118 (29.6)
Female	280 (70.4)
Age (year), mean (SD)	64.0 (8.5)
Country	
France	95 (23.9)
United States	69 (17.3)
Australia	227 (57.0)
Other	7 (1.8)
Joint affected by OA	
Knee	345 (86.7)
Hip	53 (13.3)
Duration of the flare (days)	
Mean (SD)	8.2 (10.1)
Flare of OA (during the last 4 weeks)	
Yes	258 (64.8)
No	140 (35.2)

Data are n (%) unless indicated.

SD= Standard deviation

Table II Floor and ceiling effects for the 19 items of the Flare-OA

Item content of Flare-OA	Floor effect* (%)	Ceiling effect** (%)
Painful		
My pain felt more severe compared to my usual pain	31.4	9.3
The pain did not disappear when I rested	24.4	11.8
My pain was more persistent than usual	30.9	9.8
My pain disrupted my sleep more than usual	37.0	9.1
Swelling		
My knee has been swollen more than usual	48.7	8.1
Stiffness		
I had difficulties bending at my joint (e.g. knee or hip) more than usual	30.1	10.9
I felt stiffness in my joint more than usual	25.4	10.6
Consequences of symptoms (sleep, concentration, activity, need for help, walking)		
My pain prevented me from doing certain activities during the day	20.8	14.3
I needed to put ice or something cold on my joint more than usual	59.1	7.8
I required more assistance than usual to be able to walk or stand (e.g. use of cane or crutches, leaning on railings)	57.7	7.6
I felt more restricted or impaired in my movements	23.2	14.1
I had to shorten my walking distance	26.8	19.0
I had more difficulty getting in or out of a car	26.6	10.9
Psychological aspects (mood, annoyance, frustration)		
I felt more depressed than usual	48.3	8.4
I felt frustrated because I was limited in my daily activities	29.0	19.3
I needed to avoid certain movements or activities to prevent my pain	17.0	20.9
I needed to rest (e.g. lie down or sit) to prevent my pain	28.5	12.3

There was nothing I could do to prevent my pain	38.6	10.2
I needed to change the way I performed daily activities (e.g. sitting instead of standing when getting dressed or preparing food) to prevent my pain	34.5	12.0

* Proportion of the sample at the maximum scale range >15%

** Proportion of the sample at the minimum scale range >15%

Table III Pearson correlation between Flare-OA score and scores for functional limitations on the HOOS, KOOS and Mini-OAKHQOL (n in parentheses indicates n of respondents)

Validity	**Mean (SD)**	**Correlation (r)**
HOOS		
Pain (n=46)	61.3 (23.5)	- 0.86
Symptoms (n=46)	61.5 (21.0)	- 0.71
Function in daily living (ADL) (n=44)	65.6 (23.0)	- 0.82
Function in sport and recreation (Sport/Rec) (n=44)	60.0 (28.9)	- 0.70
Quality of life (QoL) (n=44)	46.3 (26.0)	- 0.72
KOOS		
Pain (n=333)	57.6 (19.3)	- 0.72
Symptoms (n=333)	54.6 (20.1)	- 0.60
Function in daily living (ADL) (n=328)	62.9 (21.1)	- 0.68
Function in sport and recreation (Sport/Rec) (n=328)	31.5 (26.1)	- 0.52
Quality of life (QoL) (n=328)	39.8 (22.9)	- 0.64
Mini-OAKHQOL		
Pain (n=368)	53.7 (25.9)	- 0.76
Physical activities (n=368)	56.2 (27.6)	- 0.73
Mental health (n=368)	74.0 (27.2)	- 0.68
Other domains		
Social support (n=367)	56.1 (28.2)	- 0.12
Social activities (n=368)	67.7 (28.7)	- 0.32
Professional activity (n=365)	72.7 (31.5)	- 0.57
Fear of dependent (n=367)	66.0 (37.7)	- 0.64
Sexual relation (n=367)	76.6 (33.9)	- 0.42

HOOS= Hip osteoarthritis outcome score, KOOS= Knee osteoarthritis outcome score, Mini-OAKHQOL= Osteoarthritis Knee and quality of life, SD= standard deviation