Seminars in Arthritis and Rheumatism 65 (2024) 152371

Contents lists available at ScienceDirect

Seminars in Arthritis and Rheumatism

journal homepage: www.elsevier.com/locate/semarthrit

Assessing domain match and feasibility of candidate instruments matching with OMERACT endorsed domains to measure flare in knee and hip osteoarthritis

F. Queiroga [a,*], S.M. Cembalo [b], J. Epstein [a], L. Maxwell [c], T. Buttel [d], C. Copenhaver [e], M. Cross [d], D. Hunter [d], L. King [f], L. Callahan [g], L. March [d], D.E. Beaton [f], F. Guillemin [a], on behalf of the Flares in OA working group

[a] Université de Lorraine, Grand Est Region, France
[b] Patient Partner, France
[c] University of Ottawa, Ontario, Canada
[d] University of Sidney, Sydney, Australia
[e] Patient Partner, United States
[f] University of Toronto, Toronto, Canada
[g] University of North Carolina, Chapel Hill, NC, United States

ARTICLE INFO

Keywords:
Flare
Lower limbs
Osteoarthritis
Content validity
OMERACT

ABSTRACT

Purpose: To evaluate the domain match (truth) and feasibility of candidate instruments assessing flare in knee and hip osteoarthritis (OA) according to the identified domains.
Material and methods: From a literature review (575 papers), instruments were selected and evaluated using the truth and feasibility elements of the OMERACT Filter 2.2. These were evaluated by 26 experts, including patients, in two Delphi survey rounds. The final selection was obtained by a vote.
Results: 44 instruments were identified. In Delphi Round 1, five instruments were selected. In Round 2, all instruments obtained at least 75 % in terms of content match with the endorsed domains and feasibility. In the final selection, the Flare-OA questionnaire obtained 100 % favorable votes.
Conclusion: Through consensus of the working group, the Flare-OA questionnaire was selected as the best candidate instrument to move into a full assessment of its measurement properties using the OMERACT Filter 2.2.

Introduction

Osteoarthritis (OA) is a common chronic condition that affects millions of people around the world [1]. The knee and hip are two of the joints most frequently affected, and people living with OA can experience intermittent exacerbations of their symptoms or 'flares' [2,3]. Following OMERACT Handbook methodology, the following definition for the target construct of flare in OA knee and/or hip, "a transient state, different from the usual state of the condition, with a duration of a few days, characterized by worsening of pain, swelling, stiffness, impact on sleep, activity, functioning, and psychological aspects that can resolve spontaneously or lead to a need to adjust therapy" was endorsed by the OMERACT 2018 meeting [4]. The OMERACT Flares in Osteoarthritis Working Group (WG) then conducted a literature review strategy, qualitative interviews with patients and health professionals, a series of surveys and international voting at the inaugural virtual OMERACT meeting in 2020 to endorse five core domains: pain, swelling, stiffness, psychological aspects, and impact of symptoms [5].

The measurement of these five domains is needed to capture the occurrence of the targeted construct of a flare in OA. To find instruments that truly reflect the construct, i.e. measuring all core domains, the OMERACT group has developed methodological steps to accomplish this task. In this study, we present the results obtained following these steps to select candidate instrument(s) to assess flares in osteoarthritis.

* Corresponding author: 9 avenue de la Forêt de Haye, BP 20199, 54505 Vandoeuvre Les Nancy, France.
E-mail address: fabiana.queiroga@univ-lorraine.fr (F. Queiroga).

https://doi.org/10.1016/j.semarthrit.2024.152371

Available online 1 February 2024

Table 1
Percentage of agreement for each instrument according to experts' evaluation for "domains match" – Delphi Survey Round 1.

Instrument	N	Pain	Swelling	Stiffness	Impact of symptoms	Psychological aspects
Intermittent and Constant Osteoarthritis Pain questionnaire (ICOAP)	28	89 %	11 %	11 %	71 %	68 %
Fremantle Knee Awareness Questionnaire (FreKAQ)	29	24 %	34 %	14 %	34 %	14 %
Hip Injury and Osteoarthritis Outcome Score (HOOS)	29	90 %	38 %	86 %	90 %	31 %
Knee Injury and Osteoarthritis Outcome Score (KOOS)	27	81 %	67 %	78 %	85 %	33 %
Short Form McGill Pain Questionnaire (SF-MPQ-2)	27	67 %	7 %	11 %	15 %	19 %
Western Ontario & McMaster Universities Osteoarthritis Index (WOMAC)	31	71 %	16 %	68 %	71 %	13 %
Animated Activity Questionnaire (AAQ)	25	12 %	8 %	16 %	72 %	8 %
Health related quality of life measurement-Euro style (EuroQoL)	28	57 %	11 %	7 %	64 %	57 %
International Knee Documentation Committee subjective knee evaluation form (IKDC)	27	59 %	59 %	44 %	70 %	15 %
Osteoarthritis Knee and Hip Quality of Life questionnaire (OAKHQOL)	30	70 %	7 %	7 %	77 %	77 %
Flare-OA Questionnaire (Flare-OA)	29	100 %	97 %	97 %	100 %	93 %
Impact Index	29	55 %	7 %	7 %	55 %	45 %
OXFORD Hip Score (OHS)	28	75 %	14 %	29 %	29 %	25 %
Patient-Reported Outcomes Measurement Information System (PROMIS)	25	52 %	24 %	28 %	56 %	48 %
Health Survey (SF36)	28	57 %	7 %	11 %	68 %	61 %

Note: N represents the number of evaluations received by each instrument.

Material and methods

Instrument identification

In this phase, the WG applied the OMERACT Handbook Filter 2.2 [6]. The starting point was an initial literature review [7] updated by a search through PubMed, Web of Science, and PsychInfo databases covering the period 2017 to 2021 [8]. The search terms applied in databases and Prisma flowchart built for this review are presented in the Appendix. The WG evaluated the content of included instruments according to the five core domains [5], and then classified them into three categories: "sure match with domains of flare in OA"; "uncertain match with domains of flare in OA"; and "sure not match with domains of flare in OA".

Instrument selection

The OMERACT Filter is made up of three pillars of evidence to ensure an instrument is fit for the purpose of use in a Core Outcome Measurement Set in clinical trials in a given disease group or field [6,9]. They are: Truth, Feasibility and Discrimination, each must be considered for the instrument analysis by answering the questions:

- Truth: Does it match with a target domain? Does numeric score make sense?
- Feasibility: Is it practical to use?
- Discrimination: Can it discriminate between groups of interest?

In order to guide the instrument selection, we followed the Master Checklist presented in OMERACT handbook [6]. The instruments retained as "sure" or "uncertain" as a match with domains of flare in OA were evaluated by the 26 experts of the WG, composed of patients (2), physicians (6), and researchers (18), in two rounds of Delphi survey using LimeSurvey. The goal was to identify among these instruments which ones best matched with the target domain (truth) and which ones were practical to use (feasibility). The final selection was decided by a vote of the WG.

Table 2
Percentage of agreement for each instrument according to experts' evaluation for feasibility – Delphi Survey Round 1.

Instrument	N	Feasibility/Easy	Instrument	N	Feasibility/Easy
ICOAP	28	89 %	IKDC	27	59 %
FreKAQ	29	79 %	OAKHQOL	30	93 %
HOOS	29	76 %	Flare-OA	29	100 %
KOOS	27	74 %	Impact Index	29	79 %
SF-MPQ-2	27	67 %	OXFORD	28	82 %
WOMAC	31	81 %	PROMIS	25	52 %
AAQ	25	36 %	SF36	28	68 %
EuroQoL	28	75 %			

Note: N represents the number of evaluations received by each instrument.

For Round 1, the definition of Flare in OA and its domains was presented on the first survey screen and then the definition of the construct in each instrument, as well as its items, on a separate screen. The experts were asked to indicate on a five-point Likert scale to what extent the content of the questions/items matched with each Flare in OA domain (relevant items, items written at a level that will be understood by the target population, response options clear and appropriate for each item etc.). To assess feasibility, we disclosed the available information about practical considerations (easy for respondents to understand, method of administration feasible for your application, costs feasible, copyright issues reasonable and manageable). The experts were asked to indicate on a five-point Likert scale to what extent the instrument was usable. For Round 2, a survey with questions about the content of the evaluated instrument guided the WG both for truth/domains match and feasibility. These results were carefully considered and presented to the whole WG. The WG deliberated issues raised by the experts and then an anonymous vote was undertaken.

In both Delphi rounds the survey was launched using the LimeSurvey platform, with three reminders for each (the full description is shown in the Supplementary Materials). In Round 1, the WG response set was flagged with green (more than 70 % agreement), amber (between 50 and 70 %) and red (less 50 %). For Round 2, a unique threshold of 70 % to make the decision of selection was set.

Results

Instrument identification

The literature search filtered 575 papers initially. After experts' analysis, 59 studies were included, and 44 instruments associated with flare in OA were identified. Most were studies about pain in knee or hip OA (35 %), cultural adaptation of a measure (33 %) or studies investigating psychometric properties of full (16 %) or short form (4 %) instruments.

When examining each of the instruments, 16 were considered outside the scope though they were used in research on flare in knee and hip OA, they assessed other constructs such as self-efficacy, anxiety, depression or were performance-based measures. The remaining 28 instruments were assessed considering the definition and domains of Flare in OA [5,6] and 13 instruments were considered as "sure not" for domains match.

Table 3
Percentage of agreement for each instrument according to experts' evaluation for "domains match" and "feasibility" – Delphi Survey Round 2 (N = 28).

Questions	Rate	ICOAP	HOOS	KOOS	Mini-OAKHQOL	Flare-OA-19
Domains Match						
1. Are the items in this instrument relevant to you and your experience?	**F**	91 %	96 %	96 %	78 %	100 %
	Y	87 %	83 %	96 %	69 %	91 %
	U	4 %	0 %	0 %	17 %	0 %
	N	4 %	13 %	0 %	9 %	9 %
2. Were there overlapping, sensitive, or embarrassing items?*	**F**	82 %	87 %	83 %	61 %	96 %
	Y	13 %	9 %	13 %	35 %	4 %
	U	4 %	4 %	9 %	4 %	0 %
	N	78 %	83 %	74 %	57 %	96 %
3. Did you feel that all the items were clear and understandable?	**F**	92 %	87 %	91 %	87 %	95 %
	Y	83 %	78 %	87 %	78 %	91 %
	U	9 %	9 %	4 %	9 %	4 %
	N	4 %	9 %	4 %	9 %	4 %
4. Could you understand what all the questions were trying to ask? If not, which items did you feel were unclear?	**F**	87 %	91 %	91 %	91 %	91 %
	Y	74 %	87 %	87 %	78 %	78 %
	U	13 %	4 %	4 %	13 %	13 %
	N	9 %	4 %	4 %	4 %	9 %
5. Were the instructions for answering the items clear?	**F**	91 %	91 %	91 %	65 %	92 %
	Y	78 %	78 %	87 %	52 %	70 %
	U	13 %	13 %	4 %	13 %	22 %
	N	4 %	4 %	4 %	30 %	9 %
6. Does the timing of the recall period seem reasonable to you (e.g. over the past week, last 24 hours) (if applicable)?	**F**	96 %	96 %	95 %	92 %	100 %
	Y	96 %	96 %	91 %	83 %	100 %
	U	0 %	0 %	4 %	9 %	0 %
	N	0 %	0 %	0 %	4 %	0 %
Feasibility						
1. Did you find that all the items were easy to read? If not, which items were not easy to read?	**F**	96 %	91 %	95 %	91 %	95 %
	Y	96 %	91 %	91 %	91 %	91 %
	U	0 %	0 %	4 %	0 %	4 %
	N	0 %	4 %	0 %	4 %	0 %
2. Was it easy enough to complete?	**F**	91 %	91 %	95 %	83 %	95 %
	Y	87 %	87 %	91 %	83 %	91 %
	U	4 %	4 %	4 %	0 %	4 %
	N	4 %	4 %	0 %	4 %	0 %
3. Did it take a reasonable amount of time to complete?	**F**	91 %	87 %	92 %	96 %	92 %
	Y	91 %	65 %	57 %	87 %	57 %
	U	0 %	22 %	35 %	9 %	35 %
	N	4 %	9 %	4 %	4 %	4 %
4. Do you think there was too much training needed before you could be able to respond to this instrument?*	**F**	91 %	78 %	87 %	87 %	87 %
	Y	4 %	17 %	9 %	9 %	9 %
	U	0 %	4 %	0 %	4 %	0 %
	N	91 %	74 %	87 %	83 %	87 %

F = Favor; Y = Yes; U = Uncertain; N = No. Favor is the sum of Yes and Uncertain, except for questions that are reversed (marked with *).
* Percentage in "Favor" for these questions is the sum for "uncertain" and "no" because both questions are in reverse direction.

Instrument selection for truth and feasibility

In Delphi Survey Round 1, the 15 remaining instruments were rated by experts. The results for "match domain" are presented in Table 1.

Six instruments presented agreement above 70 % in at least two domains of Flare: ICOAP [10], HOOS [11], KOOS [12], WOMAC [13], OAKHQOL [14] and Flare-OA [15]. Several instruments presented a high level of agreement (between 50 and 69 %) on Pain and Impact of symptoms. The results of the evaluation of instrument feasibility are presented in Table 2.

In general, the instruments were considered easy to use. The comments provided by the experts regarding each instrument were also considered. For example, the lack of agreement of the items with the definition of flare was recurrent. The experts also commented that the psychological dimension was the least addressed domain among the instruments. The patients (the target public for the instruments) had answers that were congruent with the other experts and did not express any particular comments about feasibility. All the free-text comments obtained are shown in the Supplementary Material (Table 1.1).

In light of these results, a second Delphi survey round was conducted with five instruments: ICOAP [10], HOOS [11], KOOS [12], Mini-OAKHQOL [16] and Flare-OA [15], substituting the full length OAKHQOL (43 items) with the shorter form Mini-OAKHQOL (20 items) for ease of use. Despite a percentage of agreement close to some of this list, the WOMAC was clearly not focused on flare and was thus discarded. For this round, six questions about domains match were presented and another four about feasibility. Table 3 presents the results.

All instruments obtained at least 75 % in questions about truth and feasibility, which did not discriminate between instruments. However, analyzing these responses together with the comments on the instruments provided meaningful information. One of the principal points was that, for all but one of the five instruments, there was a lack of flare measurement and instead measured pain. The exception was the Flare-OA-19 questionnaire that received positive comments about measuring flare in agreement with its definition and dimensions. Within this phase, it was also observed that the patients' answers were well in line with the experts' answers in general, i.e. there were no overall differences to be highlighted. All the comments obtained for each instrument are shown in Supplementary Material (Table 2.1).

There were also various opinions about the ideal period to ask about flare (one versus four weeks) as well as the size of instruments (HOOS and KOOS were noted as too long), time to fill the forms, etc. The Flare-OA questionnaire obtained 100 % of the votes in favor and thus was selected as our candidate instrument.

Discussion

From two rounds of Delphi Survey, we identified a number of candidate self-report instruments providing the patient perspective on OA pain, but most did not include all the central aspects (core domains) of flare in OA endorsed by OMERACT patients and other stakeholders. Only one instrument, the Flare-OA questionnaire, met the truth and feasibility requirements for measuring flare in OA and was selected unanimously. The literature review [5] revealed that measures directed to patients with knee and hip OA frequently consider different aspects of the pain dimension and discomfort in performing daily activities, but not other dimensions. The *change of state*, a typical characteristic of flare, was only covered by the Flare-OA questionnaire [13]. Moreover, this latter was developed specifically by some authors of the present paper to match the OMERACT endorsed domains.

The rounds of Delphi Survey and the subsequent discussion on the issues raised by the experts have provided a detailed examination of all potentially useful instruments to measure flares according to the consensus definition. It was clear that the different profiles of experts were covered by the results of the evaluations. We were able to incorporate the patients' view, which is crucial to obtain a decision on feasibility. Establishing new ways to efficiently study flares will be important for patient care, especially as flares commonly occur outside the clinical setting. It means that identifying self-report measures that focus on the patient's perception could improve the accuracy of data on OA flares.

Conclusion

The OMERACT Handbook Filter 2.2 has proven to be a helpful tool for selecting a candidate instrument to measure flare in OA from a number of potential instruments. The work carried out so far addressed the OMERACT pillars of truth/domain match and feasibility. On this basis, the Flare-OA questionnaire was selected as the best matching and practical instrument to assess flare in knee and hip osteoarthritis. The next step will be to assess additional measurement properties, including responsiveness, of the Flare-OA questionnaire using the OMERACT Filter 2.2.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: none competing interests

Acknowledgement

We thank all the participants who contributed to the different stages of the work, online and on-site.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.semarthrit.2024.152371.

References

[1] Steinmetz JD, Culbreth GT, Haile LM, et al. Global, regional, and national burden of osteoarthritis, 1990–2020 and projections to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet Rheumatol 2023;5(9):e508–22. https://doi.org/10.1016/S2665-9913(23)00163-7.

[2] Thomas MJ, Guillemin F, Neogi T. Osteoarthritis flares. Clin Geriatr Med 2022;38 (2):239–57. https://doi.org/10.1016/j.cger.2021.11.001.

[3] Hawker GA. Osteoarthritis is a serious disease. Clin Exp Rheumatol 2019;37:3–6. Suppl 120(5).

[4] 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. https://doi.org/10.3899/jrheum.181085.

[5] King LK, Epstein J, Cross M, et al. Endorsement of the domains of knee and hip osteoarthritis (OA) flare: A report from the OMERACT 2020 inaugural virtual consensus vote from the flares in OA working group. Semin Arthritis Rheum 2021; 51(3):618–22. https://doi.org/10.1016/j.semarthrit.2021.03.010.

[6] Beaton D., Maxwell L., Grosskleg S., et al. Chapter 5: Instrument selection for Core Outcome Measurement Sets. In: *The OMERACT Handbook*. Filter 2.1 Updated April 1st 2021. The OMERACT Handbook © 2021 by OMERACT is licensed under CC BY-ND 4.0; 2021. Available from: https://omeracthandbook.org.

[7] 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):1920–7. https://doi.org/10.3899/jrheum.161107.

[8] Queiroga F, Cross M, Thomas JM, March L, Epstein J, Guillemin F. A systematic review of patient self-report measures of flare in knee and hip osteoarthritis: a report from the OMERACT OA Flare Working Group. Semin Arthritis Rheum 2023. Published onlineaccepted article.

[9] Maxwell LJ, Beaton DE, Boers M, et al. The evolution of instrument selection for inclusion in core outcome sets at OMERACT: Filter 2.2. Semin Arthritis Rheum 2021;51(6):1320–30. https://doi.org/10.1016/j.semarthrit.2021.08.011.

[10] Maillefert JF, Kloppenburg M, Fernandes L, et al. Multi-language translation and cross-cultural adaptation of the OARSI/OMERACT measure of intermittent and constant osteoarthritis pain (ICOAP). Osteoarthr Cartil 2009;17(10):1293–6. https://doi.org/10.1016/j.joca.2009.04.003.

[11] Nilsdotter AK, Lohmander LS, Klässbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS) – validity and responsiveness in total hip replacement. BMC Musculoskelet Disord 2003;4(1):10. https://doi.org/10.1186/1471-2474-4-10.

[12] Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee injury and osteoarthritis outcome score (KOOS)—Development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998;28(2):88–96. https://doi.org/10.2519/jospt.1998.28.2.88.

[13] 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.

[14] Rat AC, Coste J, Pouchot J, et al. OAKHQOL: A new instrument to measure quality of life in knee and hip osteoarthritis. J Clin Epidemiol 2005;58(1):47–55. https://doi.org/10.1016/j.jclinepi.2004.04.011.

[15] Traore Y, Epstein J, Spitz E, et al. Development and validation of the Flare-OA questionnaire for measuring flare in knee and hip osteoarthritis. Osteoarthr Cartil 2022;30(5):689–96. https://doi.org/10.1016/j.joca.2021.12.011.

[16] Guillemin F, Rat AC, 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. https://doi.org/10.1016/j.jclinepi.2015.06.010.