Therapeutic Innovation & Regulatory Science
https://doi.org/10.1007/s43441-023-00510-8

DIA

ORIGINAL RESEARCH

Adaptation of the WOMAC for Use in a Patient Preference Study

Sarah Stothers Rosenberg[1] · Xinyi Ng[1] · Carol Mansfield[2] · Christine Poulos[2] · Holly Peay[3] · Ting-Hsuan Lee[1] · Telba Irony[1] · Martin Ho[1]

Received: 15 August 2022 / Accepted: 7 March 2023

Abstract

Objectives To adapt a patient-reported outcome (PRO) measure, the Western Ontario McMaster Universities Osteoarthritis Index (WOMAC), into efficacy attributes for a discrete choice experiment (DCE) survey designed to quantify the relative importance of endpoints commonly used in knee osteoarthritis (KOA) trials.

Methods The adaptation comprised four steps: (1) selecting domains of interest; (2) determining presentation and framing of selected attributes; (3) determining attribute levels; and (4) developing choice tasks. This process involved input from multiple stakeholders, including regulators, health preference researchers, and patients. Pretesting was conducted to evaluate if patients comprehended the adapted survey attributes and could make trade-offs among them.

Results The WOMAC pain and function domains were selected for adaption to two efficacy attributes. Two versions of the discrete choice experiment (DCE) instrument were created to compare efficacy using (1) total domain scores and (2) item scores for "walking on a flat surface." Both attributes were presented as improvement from baseline scores by levels of 0%, 30%, 50%, and 100%. Twenty-six participants were interviewed in a pretest of the instrument (average age 60 years; 58% female; 62% had KOA for ≥ 5 years). The participants found both versions of attributes meaningful and relevant for treatment decision-making. They demonstrated willingness and ability to tradeoff improvements in pain and function separately, though many perceived them as inter-related.

Conclusions This study adds to the growing literature regarding adapting PRO measures for patient preference studies. Such adaptation is important for designing a preference study that can incorporate a clinical trial's outcomes with PRO endpoints.

Keywords Choice behavior · Patient-reported outcome measures · Osteoarthritis · Knee · Patient preference information · Discrete choice experiment or stated preference

Introduction

Eliciting quantitative patient preference information (PPI) is a scientific way to incorporate patient voices into the drug development decision-making process, and there is increasing interest in the use of PPI to support regulatory benefit-risk assessment [1–4]. Quantitative PPI measures how patients value the outcomes and endpoints of treatments, and which tradeoffs between treatment benefits and risks they consider acceptable [5]. Quantitative PPI can provide estimates on the maximum acceptable risks, minimum acceptable benefits, and relative importance of treatment outcomes and endpoints from the patients' perspectives. Discrete choice experiment (DCE) is a methodology commonly used to elicit quantitative PPI [6]. A DCE survey instrument presents respondents with a series of questions, i.e., choice tasks, each comparing two or more benefit-risk

This study was initiated and implemented when Telba Irony and Martin Ho were employees with the United States Food and Drug Administration, Center for Biologics Evaluation and Research. Telba Irony is currently working at the Janssen Pharmaceuticals Companies of Johnson and Johnson, Raritan, NJ, US. Martin Ho is currently working at Google, Palo Alto, CA, US.

This article reflects the views of the authors and should not be construed to represent FDA's views or policies.

✉ Xinyi Ng
Xinyi.Ng@fda.hhs.gov

[1] Office of Biostatistics and Pharmacovigilance, Center for Biologics Evaluation and Research, US Food and Drug Administration, Silver Spring, MD, USA

[2] RTI Health Solutions, Research Triangle Park, NC, USA

[3] RTI International, Research Triangle Park, NC, USA

Published online: 15 April 2023

profiles that are characterized by attributes (i.e., treatment features such as benefits and risks) with levels that are varied by a prespecified design. Respondents' selections across the choice tasks reflect the relative importance of changes in the levels of the benefit and risk attributes. DCEs have been widely employed in recent decades to quantify patients' and physicians' treatment preferences and the tradeoffs they are willing to accept between the benefits and risks of those treatments in a range of therapeutic areas. [7, 8]

Osteoarthritis (OA) is the most common joint disease worldwide and most frequently affects the joints of the knee (KOA) [9, 10]. The symptoms of KOA impact an estimated 14 million people in the USA, leading to functional impairment and disability that can be measured using a patient-reported outcome (PRO) measure, the Western Ontario McMaster Universities Osteoarthritis Index (WOMAC) [11–14]. To assess how patients with KOA feel and function in response to treatment, the WOMAC is often used in clinical trials to capture patient-reported pain, physical function, and stiffness [14–17]. These three domains are measured with 5 pain items, 17 function items, and 2 stiffness items. [18]

It is important to understand how patients value an outcome relative to other outcomes, as treatments may perform differently across multiple outcomes. Using the WOMAC as a case example, this study illustrates the steps and considerations involved in adapting two WOMAC domains, i.e., pain and physical function, into attributes and levels for a DCE study. Appropriate adaption of a PRO measure into attributes can be important when a clinical trial's key endpoints are based on PRO measures, and there is interest in contextualizing the preference study results with the trial's results, for example, to support regulatory decision-making. This case study serves to contribute to the growing evidence of how PRO measures may be adapted into a patient preference study. Results from the DCE study data analysis will be reported separately.

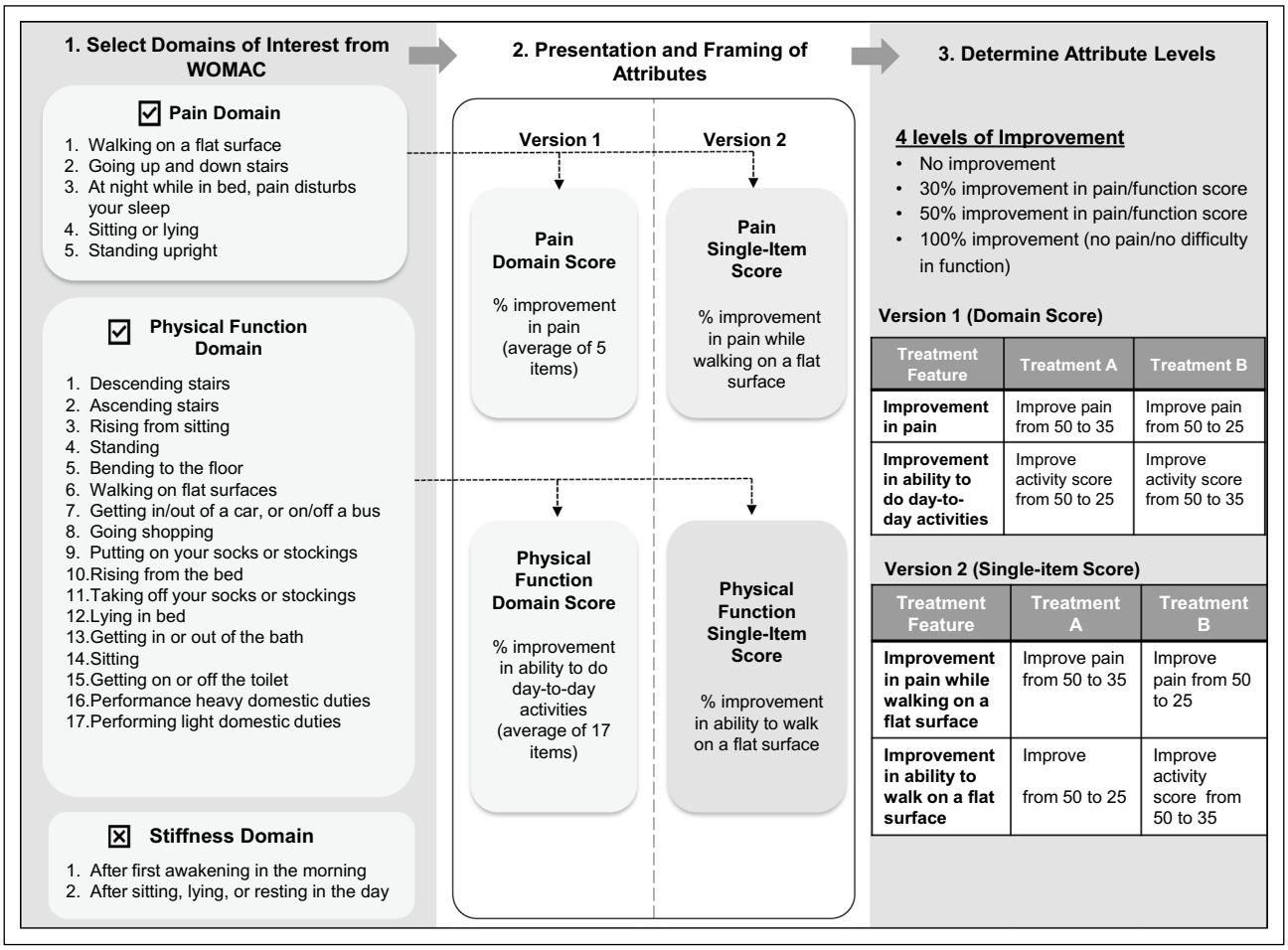


Fig. 1 Adaptation of the WOMAC domains into benefit attributes for discrete choice experiment

Methods

The three steps used for adapting the WOMAC domains into attributes are summarized in Fig. 1. These steps address the following key considerations: which domain- or item-specific scores are more appropriate for adaptation, how to best present and convey the change in WOMAC scores in the instrument, and if patients are able to weigh these domains separately and make trade-offs among them. For each step, we describe the key considerations below.

(I) Select Domains of Interest from WOMAC

The WOMAC consists of three domains: pain, function, and stiffness [18, 19]. For the purposes of this study, only two domains from the WOMAC were selected for adaptation. The two selected domains, pain and physical function, were chosen because they are commonly used as primary or secondary endpoints in recent clinical trials for KOA therapies [20–22]. The stiffness domain was less frequently used and thus, excluded from the PPI study [19]. Further, the measurement properties of the stiffness subscale are not as well-demonstrated as compared to the pain and function subscales. [18]

(II) Presentation and Framing of Attributes

The adaptation of the WOMAC pain and function domains presented a unique challenge. In previous studies, the pain and function domains were observed to be correlated [15, 23]. However, emerging therapies in KOA have different mechanisms of action and can therefore potentially lead to different levels of pain and functional improvement [22, 24]. Hence, there is interest in understanding the relative value that patients place on pain and function when they are presented with an opportunity to view improvements in these two domains separately, and the tradeoffs that patients are willing to make between the two based on their different levels of improvement. The feasibility of this was assessed in pretest interviews, as described in the "Survey Pretesting" section.

Treatment-related efficacy attributes in a DCE are often expressed as an improvement in benefit. For the selected WOMAC domains, one possible option for presentation is to describe improvements in the pain and function domain scores. However, a concern with such a presentation is that different respondents may attribute the improved domain scores to different specific aspects (or, items). For example, the pain domain score is based on scores of its 5 items (Fig. 1); if a patient selects a profile with an improvement in the pain domain score, we cannot discern if the patient is thinking of an alleviation of pain while climbing stairs or pain at night that disturbs his/her sleep, or both, or some combination of changes in the other 3 items. It is particularly important for response efficiency in PPI studies to ensure that respondents interpret the attribute in a similar fashion [25]. Presenting an improvement in domain score can potentially result in varying interpretations of the same benefit by the different respondents while they make their choices in the DCE, which could contribute to measurement error.

An improvement in a specific item score is an alternative to formulating the attribute as an improvement in a domain score. We created two versions of the DCE to explore whether results are comparable using either the WOMAC domain scores or single-item scores for pain and function.

Because the WOMAC pain and function domain scores are aggregations of multiple different item scores, identifying a single item to represent the whole domain was challenging as the selected item may not be the most relevant or salient to the respondent. A patient advisory group (PAG) of 5 patient advisors with KOA, and our partnering patient advocacy organization, the Global Healthy Living Foundation (GHLF) provided invaluable inputs. Discussions held with the PAG allowed the research group to identify single items from the pain and function domains that patients described as highly meaningful and relevant across moderate and severe KOA. Based on input from the FDA subject matter experts, climbing stairs (pain domain: "going up and down stairs"; function domain: "ascending stairs" and "descending stairs") were considered the most relevant items. However, members of the PAG reflected that people with moderate to severe KOA can make lifestyle adaptations to reduce or avoid stair climbing, while lifestyle modifications are less effective at minimizing need to walk on flat surfaces. We ultimately selected "walking on a flat surface" for both pain and function domains as this item is likely to be salient for more patients. The item also appears in both domains of the WOMAC. Using the same activity from both domains in the DCE, we can discern the value that patients place on pain relative to function improvement with regards to performing the same activity. Hence, in the item-version DCE, we used "improvement in pain while walking on a flat surface" and "improvement in ability to walk on a flat surface" to represent the improvement in pain and function attributes, respectively.

(III) Determine Attribute Levels

In a DCE, each attribute takes on various levels and differences in the levels should be meaningful to patients to facilitate trade-offs between attributes [26, 27]. The range of the attribute levels should also cover the relevant clinical values, such that the PPI results can be mapped to clinical data to inform decision making. Four levels for both the pain and function attributes were included to convey the relevant possibilities in pain and functional improvement for KOA therapies: no improvement, 30% improvement, 50% improvement or 100% improvement. Input from the FDA subject matter experts guided the selection of these levels: no improvement accounts for no changes in treatment benefit, and 30% is viewed as a meaningful improvement; 100% is used to account for future potential curative therapies; and 50% is a reasonable mid-point between 30 and 100% that would allow the researchers to detect trade-offs between different extents of improvements.

One consideration with the choice of improvement levels for an attribute was whether patients were able to distinguish the different extents of improvements presented to them. Therefore, while the levels are specified as percentage improvements from baseline, the DCE questions described the percentage improvements in terms of the corresponding change in absolute scores, which were calculated based on each respondent's baseline WOMAC pain and function scores. In this study, respondents were first asked to rate the items in both the WOMAC pain and physical domains based on a 7-day recall period using a visual analogue scale (VAS) that ranged from 0 to 100 where 0 represents no pain (or difficulty) and 100 represents extreme pain (or difficulty). Respondents were then shown their baseline pain and function score calculated as the average score across the items in the domain (domain version) or their item-specific score for pain and difficulty experienced while walking on a flat surface (item version). Before the DCE questions, descriptions of each attribute were provided, and a table was included that presented the different possible levels of improvement in pain or function based on each respondent's baseline (presented as both percentage and absolute changes in scores). For example, if a respondent reported an average score of 50 out of 100 on the 17-item function domain at baseline, a table describing the four levels of respondent-specific function score improvement from baseline was shown. In this example, a 50% improvement would also describe that the absolute function score would improve from 50 to 25.

Survey Pretesting

Pretesting is recommended by the FDA PPI guidance [1] and established good research practices by recognized professional organizations [28]. The primary aims of the pretests were to ensure that: (1) instructions were clear for patients to self-complete the survey, (2) attributes descriptions were comprehensible and relevant, and (3) respondents understood the choice questions and were able to make tradeoffs among the presented attribute levels. The pretest interviews were conducted using a "think aloud" technique, in which participants were encouraged to verbalize their thought process as they completed the survey. An experienced interviewer used probes as necessary to identify and understand any areas of confusion. The pretest interviews consisted of an iterative process where issues identified in earlier interviews were addressed by making survey revisions, and any survey changes were then tested in the subsequent interviews.

We pretested both the domain and item versions of the DCE. For the domain version, "improvement in pain" and "improvement in ability to do day-to-day activities" were used as efficacy attributes in the DCE choice tasks, whereas "improvement in pain while walking on a flat surface" and "improvement in ability to walk on a flat surface" were used in the item version. Half of the participants were randomized to answer the domain version of the survey and the other half the item version. In the pretest version of the survey, eight DCE choice tasks were included. Each choice task presented a pair of hypothetical treatments and a "continue with current treatment" profile to provide respondents the opportunity to select their status quo if they did not like either of the presented profiles.

The survey consisted of the four components: (1) respondent clinical characteristics including baseline WOMAC pain and function scores, (2) attribute descriptions and comprehension questions, (3) DCE choice tasks and (4) respondent demographic characteristics. Eligible participants for pretesting of the survey were recruited from patient panels and advertisements. They had to be 50 years or older and had a physician-confirmed or self-reported diagnosis of KOA in one or both knees. However, those who reported having axial spondyloarthropathy, gout, Lyme disease, lupus, psoriatic arthritis, or rheumatoid arthritis or if they had had knee replacement surgery in one knee (if respondent had KOA in only one knee) or both knees (if respondent had KOA in both knees) were excluded.

Table 1 Pretest participant characteristics and demographics

Characteristic	No. (%) of respondents ($N=26$)
Age (years)	
Mean (SD)	60.0 (6.9)
Range	50, 75
Gender	
Male	11(42.3%)
Female	15 (57.7%)
Race/Ethnicity[a,b]	
White	16 (66.7%)
Hispanic or Latino	0 (0.0%)
Black or African American	6 (25.0%)
Native American or American Indian	1 (4.2%)
Asian/Pacific Islander	1 (4.2%)
Other	1 (4.2%)
Educational attainment[c]	
High school or equivalent (e.g., GED)	2 (8%)
Some college but no degree	1 (4%)
Associate degree (2-year college degree)	9 (36.0%)
4-year college degree (e.g., BA, BS)	6 (24.0%)
Some graduate school but no degree	3 (12.0%)
Graduate or professional degree (e.g., MBA, MS, MD, PhD)	4 (16.0%)
Employment status	
Employed	14 (56.0%)
Retired	6 (24.0%)
Unable to work because of your osteoarthritis	2 (8.0%)
Unable to work because of disability or other health problem	3 (12.0%)
Unemployed but looking for work	1 (4.0%)
Health Insurance status[a,b]	
I do not have health insurance	0 (0.0%)
Private insurance	17 (70.8%)
Public (Medicaid, Medicare, Veteran's Health insurance)	12 (50.0%)
Household income[d]	
Less than $50,000	3 (13.0%)
$50,000 to $99,999	9 (39.1%)
$100,000 or more	6 (26.1%)
Duration from when first diagnosed, years[e]	
Less than 1 year	0 (0.0%)
At least 1 year, but less than 2-years	2 (7.7%)
At least 2-years, but less than 5-years	7 (26.9%)
At least 5-years, but less than 10-years	8 (30.8%)
At least 10-years, but less than 20-years	6 (23.1%)
More than 20-years	2 (7.7%)

SD standard deviation

[a]Two participants did not answer this question due to interview time constraints

[b]This question allows for multiple responses; therefore, the total number of responses may not add up to the total number of people who answered the question

[c]One participant did not answer this question because of interview time constraints

[d]Three participants did not answer this question and five selected "don't know or not sure" or "prefer not to say"

[e]One participant selected "don't know or not sure"

Results

Twenty-six participants were included in the pretest. Participants were on average 60 (range 50–75) years old, and 58% were female. Thirty-eight percent ($n=10$) of the respondents had physician-confirmed KOA, and 62% ($n=16$) had self-reported KOA. Table 1 summarizes their demographic and clinical characteristics.

Overall, the participants found the survey instructions clear, the choice tasks comprehensible, and they were able to understand the information and complete the survey without difficulty. Almost all respondents indicated that improvements in their KOA would be meaningful to their lives. Furthermore, participants endorsed the WOMAC attributes provided in the survey as meaningful to them, and for most, would have an impact on their treatment decision-making. This was true of both versions of the survey.

All participants were queried about the plausibility of separating pain and function. A small number of participants thought it was unrealistic that a treatment may improve function but not improve pain (and vice versa), but most participants thought that while the symptoms were related, it was plausible to separate the two. When answering the DCE choice tasks, several participants (including some participants who stated that it was realistic for pain and function to change to different degrees and some participants who thought it was unrealistic) struggled to think of these changes as being independent. However, even those that found it less realistic were still willing and able to make choices based on their preference for each attribute. Some participants were willing to choose treatments with lower functional improvement to get a treatment with greater pain improvements; this was dependent on their relative baseline pain and function levels and dependent on their level of activity. Some participants clearly favored pain improvements and others favored functional improvements.

Structure of Survey

Initially, respondents were asked to complete the pain and physical function domains of the WOMAC scale in the first section of the survey. The calculated scores were then used later in the survey to present the amount of treatment-related improvement in pain and function scores in the hypothetical profiles of the DCE survey instrument. However, during the first few pretest interviews, respondents appeared to have difficulty remembering their responses to the baseline WOMAC questions and recognizing that the scores' improvements presented in the DCE profiles were from the baseline levels they had previously provided in the survey. To improve the flow and respondent comprehension of the survey, we revised it such that the WOMAC pain and function items were presented (along with their baseline scores) with the attribute descriptions to the respondent prior to the DCE choice questions. This revision was tested in the subsequent interviews and shown to help respondents to recognize and relate their baseline and improved scores in the DCE.

Improvement With WOMAC Domains Attributes

During the pretests, a presented hypothetical improvement by 30% or 50% from baseline was not large enough for respondents who had low baseline pain scores (either domain or item score) to make tradeoffs with the other attributes. Therefore, the research team revised the eligibility criteria for the full study, restricting eligibility to respondents who rated their WOMAC item-specific pain score for walking on a flat surface 40 and above (out of 100). Many ongoing trials listed on ClinicalTrials.gov at the time of the survey development had an eligibility criterion of 40 and above for WOMAC pain domain scores. Thus, this eligibility criterion will identify a study sample that is similar to the indicated population of future potential medical products' regulatory evaluations. Respondents also had to have a baseline function score greater than 0 to allow for potential improvements in the presented hypothetical treatments.

The survey was revised based on the feedback solicited in the pretest interviews. Figure 2 presents an example of the final DCE question used in the final survey implementation. Of note, the subsequent DCE study included assessment of two other attributes relevant to treatment decision making of novel KOA therapies (e.g., cellular and tissue therapies). These two attributes included duration of improvement (patient-facing label "how long improvement lasts") and risk of tissue overgrowth (patient-facing label "risk of developing too much tissue in the knee") (Fig. 2). The duration of improvement attribute supplements information from the two WOMAC efficacy attributes in the survey as it allows us to examine if KOA patients place differential values on the duration of improvement given the different extent of improvement. Feedback from the pretesting also confirmed that duration of improvement is an important and relevant consideration for KOA patients.

Discussion

Given the growing interest in the use of PPI for decision-making, this study demonstrates via a case example in KOA, a process for adapting a PRO measure into attributes for a DCE study designed to elicit the relative importance that patients with KOA place on the two commonly used PRO domains, pain and physical function. We presented the key

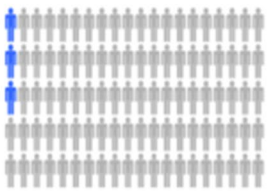
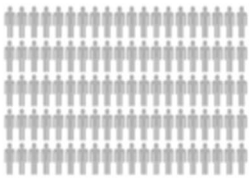
Treatment Feature	Treatment A	Treatment B	No new treatment, continue with current treatment
[Domain Score Version] Improvement in pain [Single Item Score Version] Improvement in pain while walking on a flat surface	Improve pain from [insert pain score] to [insert 30% improvement in pain score = pain score*0.70]	Improve pain from [insert pain score] to 0 (no pain)	[Domain Score Version: No additional improvement in pain] [Single Item Version: No additional improvement in pain while walking on a flat surface]
[Domain Score Version] Improvement in ability to do day-to-day activities [Single Item Score Version] Improvement in ability to walk on a flat surface	Improve activity score from [insert activity score] to 0 (no difficulty)]	Improve activity score from [insert activity score] to [insert 30% improvement in activity score = activity score*0.70]]	[Domain Score Version: No additional improvement in ability to do day-to-day activities] [Single Item Version: No additional improvement in ability to walk on a flat surface]
How long improvements last	5 years	6 months	No additional time
Risk of developing too much tissue inside the knee	 **3 out of 100 people (3%)**	 **No risk**	 **No risk**
Which treatment would you choose?	☐	☐	☐

Fig. 2 Example of a choice question used in a discrete choice experiment in knee osteoarthritis *Note* Fig. 2 is an example of a final choice question in the DCE based on the final list of attributes and levels

challenges and shared considerations encountered when developing attributes based on two domains of WOMAC.

A recent systematic review of patient preference studies in OA identified 16 studies that examined preferences for pharmaceutical, non-pharmaceutical, and surgical treatments [8]. Of the studies that focused on patient preferences for pharmacologic treatments, some assessed overall treatment benefits [29–31] while others focused on specific benefits, including pain [32–36], stiffness [33, 36] and function [32–34, 36]. In two studies whose benefit attributes were designed to correspond to the three WOMAC domains, their attribute levels were specified as absolute levels (none (0 out of 100 mm [mm]), mild (25 out of 100 mm), moderate (50 out of 100 mm), and severe (75 out of 100 mm)) instead of improvement from baseline [33, 36]. While this approach allows for the estimation of relative importance placed on these four absolute levels of the three WOMAC domains and the differences between the presented levels, it does not provide direct information regarding the values that patients place on improvement of pain and function from

baseline, unless we assume one of the presented levels to be the baseline.

Given that change from baseline WOMAC pain and function scores are endpoints frequently assessed in OA trials, our study demonstrates an approach to elicit the value patients place on these endpoints. We accomplished this by measuring the patients' baseline pain and function scores and described the improvements in the pain and function as improvements from the patient's own baseline scores. An option to remain on current treatment was included to allow patients to remain at their baseline (e.g., not accept a new treatment with additional improvement to pain and/or function). This approach allowed one to estimate the utility associated with remaining on current treatment, and the value that patients place on different degrees of pain and functional improvement from their current baseline status. Therefore, we can directly compare and consider the preference results in the context of the clinical trials' observed improvement in pain and function scores from baseline, and better gain insights on the value that patients place on these observed improvements. Being able to draw parallels between the trial and preference study results is potentially useful for regulatory decision-making.

Another important outcome from this adaptation process is the development of two different versions of DCE-formatted questions that frame the benefit attributes of pain improvement and function improvement using the following approaches: (1) overall improvement in domain scores which comprise multiple items, and (2) improvement in scores for an item (e.g., "walking on a flat surface"). One challenge with using multi-item domain scores from a PRO in a DCE is the subjective and inconsistent interpretation of improvement among the participants answering the choice questions. Hence, implementation of both versions of the DCE in the subsequent DCE study (results not reported here) will allow us to compare the consistency of results obtained from both versions and investigate reasons that led to discrepancies, if any. This will in turn provide deeper insights regarding the use of domain or item-specific scores as attributes in a DCE.

We also described how patients' input was sought in the adaptation process. First, we sought feedback from patient advisors and partnered with a patient advocacy organization to identify the most relevant item in the WOMAC pain and function domains to use in our item-version of the DCE. Second, we pretested the survey with patients through cognitive interviews and revised several components of the survey based on the feedback. The importance of involving patients in preference studies is well-recognized, and we demonstrated how patient input was incorporated into the development of the survey instrument.

Most of the reviewed studies examining pharmacologic treatment preferences did not account for the duration of improvement in their assessment except Copsey et al. [36] This study observed that duration of improvement can play an important role in patients' treatment decision-making since KOA is a chronic and debilitating condition. Estimating the value that patients place on duration of improvement is useful in allowing us to evaluate the observed improvements in WOMAC scores in the trials at different timepoints. Furthermore, including a duration attribute also allows us to examine the interaction between extent and duration of improvement. Specifically, based on the subsequent preference elicitation study using the developed DCE, we will be able to examine if KOA patients place differential values on the duration of improvement given the different extent of improvement (e.g., patients might value a short duration [e.g., 6-months] of a large improvement from baseline [e.g., 100%] much more than a short duration [e.g., 6-months] of a small improvement from baseline [e.g., 30%]).

There are some limitations to note in this study. First, this is a case study demonstrating the adaptation of a specific PRO measure (i.e., WOMAC) for a PPI study and may not be generalizable to all contexts. Developing a recommended approach to adapting PROMs to PPI studies would be highly useful for PPI research. Future research is needed to explore if this approach can be applied to or adapted for other instruments or therapeutic areas. The process of ultimately selecting and adapting a PRO measure for use in a DCE may differ depending on the pre-specified condition, the research question, the context of use, and the availability of a PRO measure. In addition, the interviews described in this manuscript were conducted during the onset of the COVID-19 pandemic (April 2020); and thus, while we intended to obtain physician-confirmed diagnosis using a doctor's note as confirmation, we were unable to do so for part of the sample and had to rely on self-reported diagnoses for some participants because physician offices were restricting non-urgent visits (10 out 26 participants in the pretest had a self-reported KOA diagnosis). However, for the subsequent elicitation study, we recruited a sample of patients with physician-confirmed diagnosis and a sample of patients with self-reported diagnosis and compared results from both sources.

In conclusion, the study provides a case example to demonstrate how a PRO measure can be adapted for use in a patient preference study among patients with KOA. The developed DCE survey was implemented in a patient preference study, and the results elicited and quantified the value that KOA patients place on the endpoints commonly used in KOA product development. Preliminary results of the DCE study are reported separately [37, 38]

Acknowledgements
The authors thank Larissa Lapteva, M.D. from FDA CBER Office of Tissues and Advanced Therapies for her review of this article.

Funding
This study is funded by the US Food and Drug Administration.

Data Availability
Data sharing is not applicable to this article as no datasets were generated or analyzed in this article.

References

1. U.S. Food and Drug Administration (FDA). Patient Preference Information – Voluntary Submission, Review in Premarket Approval Applications, Humanitarian Device Exemption Applications, and De Novo Requests, and Inclusion in Decision Summaries and Device Labeling Guidance for Industry, Food and Drug Administration Staff, and Other Stakeholders. Accessed April 20, 2021. https://www.fda.gov/media/92593/download
2. Ho M, Saha A, McCleary KK, et al. A framework for incorporating patient preferences regarding benefits and risks into regulatory assessment of medical technologies. Value Health. 2016;19(6):746–50. https://doi.org/10.1016/j.jval.2016.02.019.
3. Ho MP, Gonzalez JM, Lerner HP, et al. Incorporating patient-preference evidence into regulatory decision making. Surg Endosc. 2015;29(10):2984–93. https://doi.org/10.1007/s00464-014-4044-2.
4. U.S. Food and Drug Administration (FDA). Benefit-Risk Assessment for New Drug and Biological Products Guidance for Industry - Draft Guidance. Accessed December 8, 2021. https://www.fda.gov/media/152544/download
5. Benz HL, Saha A, Tarver ME. Integrating the voice of the patient into the medical device regulatory process using patient preference information. Value Health. 2020;23(3):294–7. https://doi.org/10.1016/j.jval.2019.12.005.
6. Soekhai V, Whichello C, Levitan B, et al. Methods for exploring and eliciting patient preferences in the medical product lifecycle: a literature review. Drug Discov Today. 2019;24(7):1324–31. https://doi.org/10.1016/j.drudis.2019.05.001.
7. Soekhai V, de Bekker-Grob EW, Ellis AR, Vass CM. Discrete choice experiments in health economics: Past Present and Future. Pharmacoeconomics. 2019;37(2):201–26. https://doi.org/10.1007/s40273-018-0734-2.
8. Al-Omari B, McMeekin P, Bate A. Systematic review of studies using conjoint analysis techniques to investigate patients' preferences regarding osteoarthritis treatment. Patient Prefer Adher. 2021;15:197–211. https://doi.org/10.2147/ppa.S287322.
9. Vina ER, Kwoh CK. Epidemiology of osteoarthritis: literature update. Curr Opin Rheumatol. 2018;30(2):160–7. https://doi.org/10.1097/bor.0000000000000479.
10. Glyn-Jones S, Palmer AJ, Agricola R, et al. Osteoarthritis. Lancet. 2015;386(9991):376–87. https://doi.org/10.1016/s0140-6736(14)60802-3.
11. 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;1:95–108.
12. Deshpande BR, Katz JN, Solomon DH, et al. Number of persons with symptomatic knee osteoarthritis in the us: impact of race and ethnicity, age, sex, and obesity. Arthritis Care Res (Hoboken). 2016;68(12):1743–50. https://doi.org/10.1002/acr.22897.
13. Fautrel B, Hilliquin P, Rozenberg S, et al. Impact of osteoarthritis: results of a nationwide survey of 10,000 patients consulting for OA. Joint Bone Spine. 2005;72(3):235–40. https://doi.org/10.1016/j.jbspin.2004.08.009.
14. Woolacott NF, Corbett MS, Rice SJC. The use and reporting of WOMAC in the assessment of the benefit of physical therapies for the pain of osteoarthritis of the knee: findings from a systematic review of clinical trials. Rheumatology. 2012;51(8):1440–6. https://doi.org/10.1093/rheumatology/kes043.
15. Lorenzoni G, Azzolina D, Soriani N, Gregori D. Evaluating therapeutic effect on WOMAC subscales in osteoarthritis RCTs: When model choice matters. J Eval Clin Pract. 2018;24(1):89–96. https://doi.org/10.1111/jep.12729.
16. Smith PA. Intra-articular autologous conditioned plasma injections provide safe and efficacious treatment for knee osteoarthritis: an FDA-sanctioned, randomized, double-blind, placebo-controlled clinical trial. Am J Sports Med. 2016;44(4):884–91. https://doi.org/10.1177/0363546515624678.
17. Dworkin RH, Peirce-Sandner S, Turk DC, et al. Outcome measures in placebo-controlled trials of osteoarthritis: responsiveness to treatment effects in the report database. Osteoarthr Cartil. 2011;19(5):483–92. https://doi.org/10.1016/j.joca.2011.02.020.
18. McConnell S, Kolopack P, Davis AM. The Western ontario and mcmaster universities osteoarthritis index (WOMAC): a review of its utility and measurement properties. Arthritis Rheum. 2001;45(5):453–61. https://doi.org/10.1002/1529-0131(200110)45:5%3c453::Aid-art365%3e3.0.Co;2-w.
19. Copsey B, Thompson JY, Vadher K, et al. Problems persist in reporting of methods and results for the WOMAC measure in hip and knee osteoarthritis trials. Qual Life Res. 2018;28(2):335–43. https://doi.org/10.1007/s11136-018-1978-1.
20. Berenbaum F, Blanco FJ, Guermazi A, et al. Subcutaneous tanezumab for osteoarthritis of the hip or knee: efficacy and safety results from a 24-week randomised phase III study with a 24-week follow-up period. Ann Rheum Dis. 2020;79(6):800–10. https://doi.org/10.1136/annrheumdis-2019-216296.
21. Kon E, Engebretsen L, Verdonk P, Nehrer S, Filardo G. Clinical outcomes of knee osteoarthritis treated with an autologous protein solution injection: A 1-year pilot double-blinded randomized controlled trial. Am J Sports Med. 2018;46(1):171–80. https://doi.org/10.1177/0363546517732734.
22. Grässel S, Muschter D. Recent advances in the treatment of osteoarthritis. F1000Res. 2020. https://doi.org/10.12688/f1000research.22115.1.
23. Parkes MJ, Callaghan MJ, Tive L, Lunt M, Felson DT. Responsiveness of single versus composite measures of pain in knee osteoarthritis. J Rheumatol. 2018;45(9):1308–15. https://doi.org/10.3899/jrheum.170928.
24. Latourte A, Kloppenburg M, Richette P. Emerging pharmaceutical therapies for osteoarthritis. Nat Rev Rheumatol. 2020;16(12):673–88. https://doi.org/10.1038/s41584-020-00518-6.
25. Reed Johnson F, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR conjoint analysis experimental design good research practices task force. Value Health. 2013;16(1):3–13. https://doi.org/10.1016/j.jval.2012.08.2223.
26. Kløjgaard ME, Bech M, Søgaard R. Designing a stated choice experiment: the value of a qualitative process. J Choice Model. 2012;5(2):1–18. https://doi.org/10.1016/s1755-5345(13)70050-2.
27. Cook NS, Cave J, Holtorf AP. Patient preference studies during early drug development: aligning stakeholders to ensure development plans meet patient needs hypothesis and theory. Front Med

(Lausanne). 2019;6(82):82. https://doi.org/10.3389/fmed.2019.00082.
28. Bridges JF, Hauber AB, Marshall D, et al. Conjoint analysis applications in health–a checklist: a report of the ISPOR good research practices for conjoint analysis task force. Value Health. 2011;14(4):403–13. https://doi.org/10.1016/j.jval.2010.11.013.
29. Fraenkel L, Bogardus ST Jr, Concato J, Wittink DR. Treatment options in knee osteoarthritis: the patient's perspective. Arch Intern Med. 2004;164(12):1299–304. https://doi.org/10.1001/archinte.164.12.1299.
30. Fraenkel L, Wittink DR, Concato J, Fried T. Informed choice and the widespread use of antiinflammatory drugs. Arthritis Rheum. 2004;51(2):210–4. https://doi.org/10.1002/art.20247.
31. Al-Omari B, Sim J, Croft P, Frisher M. Generating individual patient preferences for the treatment of osteoarthritis using adaptive choice-based conjoint (ACBC) Analysis. Rheumatol Ther. 2017;4(1):167–82. https://doi.org/10.1007/s40744-017-0056-4.
32. Ratcliffe J, Buxton M, McGarry T, Sheldon R, Chancellor J. Patients' preferences for characteristics associated with treatments for osteoarthritis. Rheumatology (Oxford). 2004;43(3):337–45. https://doi.org/10.1093/rheumatology/keh038.
33. Hauber AB, Arden NK, Mohamed AF, et al. A discrete-choice experiment of United Kingdom patients' willingness to risk adverse events for improved function and pain control in osteoarthritis. Osteoarthr Cartil. 2013;21(2):289–97. https://doi.org/10.1016/j.joca.2012.11.007.
34. Hiligsmann M, Dennison E, Beaudart C, et al. A discrete-choice experiment to assess patients' preferences for osteoarthritis treatment: An ESCEO working group. Semin Arthritis Rheum. 2020;50(5):859–66. https://doi.org/10.1016/j.semarthrit.2020.08.005.
35. Laba TL, Brien JA, Fransen M, Jan S. Patient preferences for adherence to treatment for osteoarthritis: the MEdication decisions in osteoarthritis study (MEDOS). BMC Musculoskelet Disord. 2013;14(1):160. https://doi.org/10.1186/1471-2474-14-160.
36. Copsey B, Buchanan J, Fitzpatrick R, Lamb SE, Dutton SJ, Cook JA. Duration of treatment effect should be considered in the design and interpretation of clinical trials: results of a discrete choice experiment. Med Decis Making. 2019;39(4):461–73. https://doi.org/10.1177/0272989x19841877.
37. Poulos C, Mansfield C, Peay H, et al. Eliciting Treatment Preferences in Knee Osteoarthritis: Using Domain versus Single Item Scores of the Western Ontario McMaster Universities (WOMAC) Osteoarthritis Index (Oral Abstract Presentation). presented at: 43rd Society of Medical Decision-Making (SMDM) 2021; Virtual.
38. Poulos C, Ng X, Mansfield C, et al. Logical soundness in a discrete choice experiment study of benefit risk preferences. ISPOR 2022 Conference (poster). Accessed May 26, 2022. https://www.rtihs.org/publications/logical-soundness-discrete-choice-experiment-study-benefit-risk-preferences