



Validation of a new symptom outcome for knee osteoarthritis: the Ambulation Adjusted Score for Knee pain

Grace H. Lo^{1,2} · Jing Song³ · Timothy E. McAlindon⁴ · Gillian A. Hawker⁵ · Jeffrey B. Driban⁴ · Lori Lyn Price⁶ · Charles B. Eaton⁷ · Marc C. Hochberg⁸ · Rebecca D. Jackson⁹ · C. Kent Kwok¹⁰ · Michael C. Nevitt¹¹ · Dorothy D. Dunlop³

Received: 22 August 2018 / Revised: 10 October 2018 / Accepted: 24 October 2018 / Published online: 12 November 2018
© International League of Associations for Rheumatology (ILAR) 2018

Abstract

Objective Validation of a symptom measure for early knee OA may help identify new treatments and modifiable risk factors. Symptom measures that consider pain in the context of activity level may provide better discrimination than pain alone. Therefore, we aimed to compare sensitivity to change for radiographic progression between Ambulation Adjusted Score for Knee pain (AASK), which accounts for self-reported ambulation, and Western Ontario McMaster Osteoarthritis (WOMAC) knee pain score.

Design Participants were assessed annually up to 48 months using WOMAC, Physical Activity Scale for the Elderly (PASE) ambulation, and knee radiographs. AASK was defined as ((WOMAC pain) + 1)/((average daily hours of walking) + 1). Radiographs were scored for Kellgren-Lawrence (KL) grade. Linear regression, stratified by OA status, evaluated relationships between changes in AASK and WOMAC pain and KL grade over time.

Results For 4191 people (8030 knees), the mean age was 61.2 (+9.2) years old and BMI was 28.6 (+4.8) kg/m²; 58% female. Over 40% of knees had WOMAC pain scores of 0; by design, no knees had AASK scores of 0. Annual changes in AASK were more sensitive to changes in KL than changes in WOMAC in those without baseline OA (0.20 and 0.16 change per unit KL change, $p=0.005$ and 0.070 respectively), but performed similarly in knees with OA.

Conclusion AASK is simple to assess using existing validated questionnaires. AASK performs well in individuals with and without OA and should be considered in clinical trials and observational studies of early knee OA.

Keywords Knee · Measurement · Osteoarthritis · Outcome measure · Validation

✉ Grace H. Lo
ghlo@bcm.edu

¹ Department of Medicine, Baylor College of Medicine, 1 Baylor Plaza, BCM-285, Houston, TX 77030, USA

² Medical Care Line and Research Care Line, Houston VA HSR&D Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey Medical Center, Houston, TX, USA

³ Center for Healthcare Studies, Institute for Public Health and Medicine, Northwestern Feinberg School of Medicine, Chicago, IL, USA

⁴ Division of Rheumatology, Tufts Medical Center, Boston, MA, USA

⁵ Department of Medicine, University of Toronto, Toronto, ON, Canada

⁶ The Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, and Tufts Clinical and Translational Science Institute, Tufts University, Boston, MA, USA

⁷ Department of Family Medicine and Epidemiology, Warren Alpert Medical School of Brown University, Providence, RI and Department of Epidemiology, School of Public Health of Brown University, Providence, RI, USA

⁸ Departments of Medicine and Epidemiology and Public Health, University of Maryland School of Medicine, Baltimore, MD, USA

⁹ Division of Endocrinology, Diabetes and Metabolism, The Ohio State University, Columbus, OH, USA

¹⁰ University of Arizona Arthritis Center, University of Arizona, Tucson, AZ, USA

¹¹ Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, CA, USA

Introduction

The public health burden of osteoarthritis (OA) is substantial, as the most common form of arthritis, with at least 30 million adults in the USA with clinical OA [1], and with evidence that its prevalence is increasing over time [2]. The direct costs related to the disease are impressive, estimated to cost \$189 billion within the USA annually [3]. Additionally, the indirect costs related to OA include 68 million work loss days per year within the USA [4] likely related to the fact that it is the most common cause of dependency in lower limb tasks [5].

The concept of a *disease modifying osteoarthritis drug* (DMOAD) was first described 24 years ago [6] with the idea that identifying a pharmacologic agent could slow the progression of articular cartilage damage within a joint. Unfortunately, thus far, no DMOADs have been identified. Additionally, most known risk factors for knee OA are difficult to modify. In a recent review of knee OA prevention, increasing age, female sex, obesity, and prior knee injuries were identified as convincing risk factors for incident knee OA [7]. While identification of risk factors is important for better understanding the pathophysiology of knee OA, two of these risk factors, age and sex, are not modifiable, while the other two, obesity and prior knee injuries, are difficult to prevent or modify. The lack of success in these endeavors may be because trials and observational studies of OA usually included only participants with established symptomatic OA. Instead, the window of opportunity to modify the natural history of OA may occur early in its development. Thus, it is critical that we have instruments that can measure symptoms among those with early disease to facilitate the identification of a credible DMOAD and potentially modifiable risk factors for the disease.

The Western Ontario McMaster Osteoarthritis score (WOMAC) [8–12] is an FDA-approved measure of knee pain and the most commonly used symptom assessment in knee OA observational studies and clinical trials [13]. The WOMAC pain scale assesses self-reported pain when performing five activities (at rest, lying, going up and down stairs, walking, and standing) but it does not adjust for the amount of activity that a person engages in at the time that the pain is assessed. Because there is a substantial floor effect of the WOMAC, particularly among those with early disease [14, 15], we propose that an adjusted pain score that accounts for participant self-reported ambulation, such as the Ambulation Adjusted Score for Knee pain (AASK), will be more sensitive than WOMAC to changes in radiographic disease status.

It has long been asserted that knee pain related to OA does not always associate well with radiographic severity [16–18] perhaps because people adjust activities to manage symptoms. Therefore, measures of symptoms that consider pain in the context of activity level may provide better discrimination than pain alone. Because people with knee OA often identify

worsening pain with activity [19], they can avoid activities that precipitate pain as a strategy to manage their pain [20]. It follows that symptom severity should be considered in the context of the individual's level of physical activity. Measures of symptoms that consider pain in the context of activity level may provide better discrimination of disease severity than pain alone, particularly among those with earlier disease.

Consider the following situation, if two people both grade their WOMAC pain to be a 5, but one person has a very sedentary lifestyle but the other runs 5 miles daily, the WOMAC pain score would classify these two individuals as being equally symptomatic. An adjusted pain score that accounts for participant self-reported ambulation would grade the sedentary individual as being more symptomatic compared to the person who runs 5 miles daily.

In this study, we aimed to compare the longitudinal relationships of AASK and WOMAC with changes in radiographic OA severity within the Osteoarthritis Initiative (OAI), a multi-center observation study, mostly comprised of people who are either at risk for or who have symptomatic knee OA.

Methods

Study design

This study is nested within the OAI, a prospective multi-center observational study of 4796 participants in three groups: progression ($N = 1389$), incidence ($N = 3285$), and a non-exposed control group ($N = 122$). “Progression sub-cohort” members have pre-existent symptomatic radiographic knee OA (ROA), “incident sub-cohort” members are at increased risk for symptomatic ROA based on the presence of known OA risk factors, and “non-exposed control sub-cohort” members have neither symptomatic ROA nor are at risk for symptomatic ROA. At the time of OAI enrollment (February 2004–May 2006), participants were men and women ages 45 to 79 years old recruited at one of four clinical sites: Memorial Hospital of Rhode Island (Pawtucket, RI), Ohio State University (Columbus, Ohio), University of Pittsburgh (Pittsburgh, PA), and University of Maryland/Johns Hopkins University (Baltimore, MD).

For the baseline analyses, we included all OAI participants who, at OAI baseline, had at least one native knee and respective radiographic severity readings, WOMAC scores, and responses to the ambulation questions from the Physical Activity Scale for the Elderly (PASE) questionnaire.

For the longitudinal analyses, we included all OAI participants who had at least one native knee and respective radiographic severity readings, WOMAC scores, and responses to the ambulation questions from the PASE. Knees with Kellgren-Lawrence grade 4 (the maximal score for

radiographic OA severity) at the baseline of the observation period were excluded from these analyses.

All publicly available data were accessed from OAI website (<http://oai.epi-ucsf.org/datarelease/>).

Pain assessment

At baseline and annual clinic follow-up visits, participants were asked to self-report knee-specific pain in reference to the last 7 days by completing the Western Ontario and McMaster (WOMAC) Universities Osteoarthritis Pain Scale (3.1 Likert version) [9, 11] separately for each knee. The scale assesses pain severity with five activities: pain with walking, taking stairs, standing, in bed, and lying down or sitting. Possible pain scores range from 0 (no pain) to 20 (severe pain).

Physical activity risk factor assessments

At baseline and annual clinic follow-up visits, OAI participants completed the PASE questionnaire which is a validated questionnaire [21]. In this questionnaire, there were two questions that specifically focused on walking. The first question was “Over the past 7 days, how often did you take a walk outside your home or yard for any reason? For example for fun or exercise, walking to work, walking the dog, etc.?” Participants could select one of 4 options: (1) Never, (2) Seldom/1–2 days, (3) Sometimes/3–4 days, or (4) Often/5–7 days. Participants who responded seldom, sometimes, or often to the prior question were asked this next question: “On average, how many hours per day did you spend walking?” The participants could select one of the following options: (1) less than 1 hour, (2) Greater than 1 hour but less than 2 hours, (3) Greater than 2 hours but less than 4 hours, or (4) 4 hours or more. Using these two questions, *average daily hours of walking* was calculated by dividing the total hours of walking over the last 7 days by 7. Based on the literature, the PASE test-retest reliability coefficient was 0.75 (95% CI=0.69–0.80) from testing in 254 people over a 3–7-week interval, indicating that this is a reliable test [21].

Definition of new symptom score

The AASK was defined as $((\text{WOMAC pain}) + 1) / ((\text{average daily hours of walking}) + 1)$. We added 1 to the numerator to include those reporting no pain and 1 to the denominator to include those reporting no walking. Based on this definition, AASK gives a symptom score distribution even when WOMAC pain = 0.

Knee radiographs

Bilateral, fixed-flexion posterior-anterior (PA) knee radiographs [22] were obtained at the OAI baseline and annual clinic visits through 48 months. Films were obtained in a standing position with knees flexed 20–30 degrees and feet rotated internally 10 degrees. A SynaFlexer plexiglass frame was used to fix the position of the knees and feet [23]. Central readers assessed OA severity, Kellgren and Lawrence (KL) grade (0–4) using the Osteoarthritis Research Society International Atlas [24]. The reliability for these readings (read-reread) was substantial [25] (kappa coefficient for intra-rater agreement ranged from 0.70–0.80) [26]. Radiographic OA was defined as KL grade ≥ 2 . As we are looking to evaluate how symptom assessments perform over the spectrum of disease, KL grade is an appropriate radiographic measure that can assess transition to incident disease as well as worsening of disease once the disease is established.

Statistical analysis

SAS version 9.4 (SAS Institute Inc., Cary, NC, USA) was used for all analyses, and $p \leq 0.05$ was considered statistically significant. Baseline analyses compared the empirical distributions of the AASK and WOMAC pain employing a Kolmogorov-Smirnov test with a jackknifed variance estimate to account for clustered data due to the inclusion of both knees. We also presented histograms of AASK and WOMAC pain stratified by KL grade.

Longitudinal analyses evaluated knee-specific change in KL score over each 12-month follow-up period to separately predict the outcomes: change in AASK and WOMAC pain scores over the same period. Observation periods where knees had a KL score of 4 at the beginning of the period were excluded. We included bilateral knees, and time periods 0 to 12 months, 12 to 24 months, 24 to 36 months, and 36 to 48 months until the first occurrence of a KL score of 4, arthroplasty, or the 48-month observation. Because visits after 48 months did not occur on an annual basis, subsequent time periods were excluded. Observation periods where knees had an arthroplasty by the end of the period were also excluded. If BMI was missing (0.2%), the most recent annual assessment was used as a proxy or the time period was excluded (0.04%). We performed linear regression with generalized estimating equations to adjust covariates (age, gender, BMI) and account for correlation between knees and across time points within a given person. We also presented stratified results by radiographic OA (ROA) status.

Results

Baseline analyses We included 4480 people who contributed 8873 knees, mean baseline age of 61.2 (SD 9.2) years old and

baseline BMI of 28.6 (SD 4.8) kg/m² with 58% female. One thousand nine hundred fifty participants did not have ROA at baseline while 2530 did (Table 1). The people with ROA were slightly older, had a greater BMI, and had higher AASK and WOMAC scores (Table 1). Their daily ambulation estimates were similar (Table 1). Figure 1 shows the empirical distribution for baseline AASK and WOMAC pain scores for all native knees. For the WOMAC, there is a stair-step appearance to the distribution while for AASK, the curve is smooth. The distributions of these curves were significantly different (Kolmogorov-Smirnov test, $p < 0.0001$). More than 40% of knees had a WOMAC pain score of 0, illustrating the floor effect exhibited by this measure. In contrast, AASK scores show a range of values among the knees with a WOMAC pain score of 0. We found that of the people who had an AASK score of less than 1, 89% had a WOMAC score of 0, confirming that AASK assists in addressing the floor effect exhibited by WOMAC pain. Figure 2 also highlights the floor effect of WOMAC across multiple grades of KL severity which is not seen with the AASK score.

Longitudinal analyses Four thousand one hundred ninety-one people were included, contributing 8030 knees (mean baseline age 61.2 (SD 9.2) years, baseline BMI 28.6 (SD 4.8) kg/m², and 58% female). Longitudinal analyses using linear regression evaluated the relationship of radiographic severity score changes with the change in each symptom score. The annual change in KL score was significantly associated with annual changes in both AASK (0.24 change per unit KL change, $p < 0.001$) and WOMAC (0.22 change per unit KL change, $p = 0.002$) scores, as shown in Table 2. When analyses were

performed stratified by ROA (KL ≥ 2 and KL < 2), annual changes in AASK and WOMAC remained significantly associated with annual changes in KL score among individuals with established ROA, but in individuals without ROA, a significant relationship was found for AASK (0.20 change per unit KL change, $p = 0.005$) but not WOMAC (0.16 change per unit KL change, $p = 0.070$). 0.03% of the AASK scores were outliers (with a value > 20) so we performed analysis with and without these observations. There was no change in the results with exclusion of these observations. Recognizing PASE is a person-level assessment, sensitivity analyses were conducted using person-level pain and K/L grade based on the worst score from both knees, which yielded almost identical findings (not shown).

Discussion

Our study comparing AASK and WOMAC pain indicates AASK can be simply measured using WOMAC [10] and the walking questions in the PASE [21] questionnaire. Both AASK and WOMAC pain are associated with change in radiographic severity among participants with established radiographic disease. However, among those initially without ROA, AASK had a stronger association than WOMAC pain to subsequent changes in radiographic severity. The empirical distributions of AASK and WOMAC pain indicate that AASK may perform better than WOMAC in the lower range of the scores because AASK addresses a substantial floor effect exhibited by the WOMAC pain (Fig. 1). Observational studies and clinical trials targeting risk factors

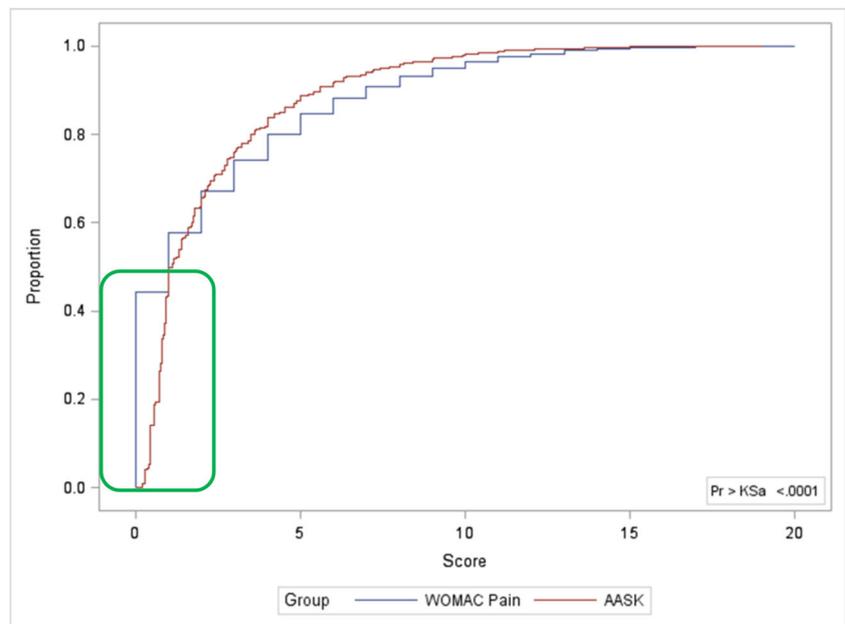
Table 1 Baseline characteristics of participants

	All participants (<i>n</i> = 4480)	No ROA (<i>n</i> = 1950)	ROA (<i>n</i> = 2530)
Age, mean (SD)	61.2 (9.2)	59.4 (9.1)	62.6 (9.0)
Sex, %			
Male	42.0	41.6	42.2
Female	58.0	58.4	57.8
BMI, mean (SD)**	28.6 (4.8)	27.2 (4.5)	29.6 (4.8)
KL grade, %*			
KL grade 0	28.1	64.5	
KL grade 1	15.5	35.5	
KL grade 2	30.2		53.5
KL grade 3	19.8		35.0
KL grade 4	6.5		11.5
WOMAC pain*, median (IQR)	2.0 (0.0–5.0)	1.0 (0.0–4.0)	3.0 (1.0–6.0)
Walking estimate (average hours walked per day), median (IQR)	0.4 (0.1–1.3)	0.4 (0.1–1.3)	0.4 (0.1–1.3)
AASK*, median (IQR)	2.0 (0.9–4.0)	1.4 (0.8–3.0)	2.4 (1.0–4.9)

*For K/L grade, WOMAC pain score, and AASK score, the higher value between the two knees was used as the person-level measure

**4476 persons for the overall cohort, 1950 for the no OA cohort, and 2526 for the OA cohort

Fig. 1 Empirical distribution for WOMAC pain and AASK scores of OAI baseline visit for all native knees. The distributions of these curves were significantly different (Kolmogorov-Smirnov test, $p < 0.0001$). Note the stepwise appearance to WOMAC pain compared to the smooth curve of the AASK. The green box highlights the severe floor effect of the WOMAC where 40% of knees have a score of 0 whereas AASK scores show a distribution of scores in this range



and treatments early in disease will only have a chance of success if instruments that measure symptoms during that time

frame within OA are available. Thus, it is salient that our findings confirm that measures of symptoms that consider

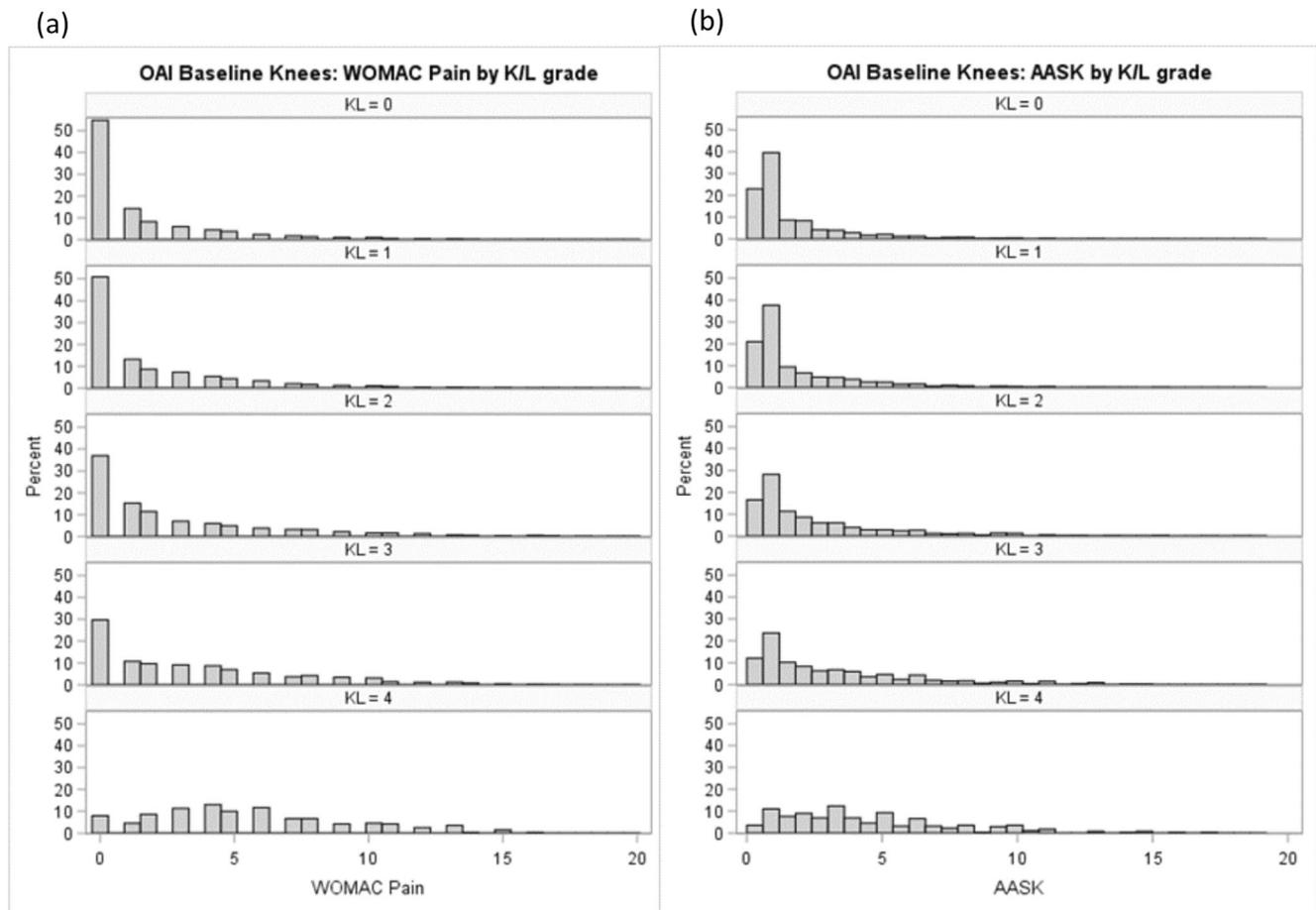


Fig. 2 Histograms of OAI baseline WOMAC (a) and AASK (b) stratified by KL scores. $AASK = ((WOMAC) + 1) / ((\text{average daily hours of walking}) + 1)$. Calculated scores greater than 20 were assigned to a score of 20

Table 2 Longitudinal analysis using linear regression: For all participants, the predictor is change in KL score and the outcome is difference in symptom scores over 1 year of follow-up. For all scores, higher score = greater symptoms

	Unadjusted change per unit change in KL score over 1 year of follow-up (95% CI); <i>p</i> value		Adjusted ^a change per unit change in KL score over 1 year of follow-up (95% CI); <i>p</i> value	
	WOMAC	AASK	WOMAC	AASK
Overall analyses				
All participants (<i>n</i> = 26,977 knee-years, 8030 knees, 4191 participants)	0.22 (0.08, 0.35); 0.002	0.24 (0.12, 0.36); < 0.001	0.22 (0.08, 0.35); 0.002	0.24 (0.12, 0.35); < 0.001
Stratified analyses				
Knees without OA (<i>n</i> = 15,579 knee-years, 4712 knees, 2908 participants)	0.16 (−0.01, 0.33); 0.068	0.20 (0.06, 0.34); 0.005	0.16 (−0.01, 0.33); 0.070	0.20 (0.06, 0.33); 0.005
Knees with OA (<i>n</i> = 11,398 knee-years, 3647 knees, 2414 participants)	0.42 (0.13, 0.70); 0.004	0.38 (0.13, 0.63); 0.003	0.42 (0.14, 0.70); 0.004	0.38 (0.13, 0.63); 0.003

^a Adjusted for age, sex, and BMI

AASK = ((WOMAC) + 1)/((average daily hours of walking) + 1)

0.03% of AASK scores were outliers with scores greater than 20. When these scores were excluded, this did not change our results

pain in the context of activity level provide better discrimination of disease severity than pain alone, particularly among those with lower levels of disease severity.

Our development of AASK is not the first attempt to target those with earlier symptoms. The Knee Injury and Osteoarthritis Outcome Score (KOOS) [27] also has had a similar aim. The KOOS does this by including assessments of higher functioning activities such as pivoting and jumping. Although this does address higher levels of activity, ultimately many people end up not answering these questions presumably for a variety of reasons including that some may have a preference to not do those activities because they are uninterested in those activities or alternatively they might avoid them to prevent symptoms. These groups of people potentially have the extremes of symptoms so scoring them both as zeros is problematic. The current recommendation by the developers of KOOS is to exclude the questions that are not answered which means that these all people who do not respond to these high levels of activity questions are treated in the same way. In contrast to KOOS, AASK adjusts for activity by accounting for a quantitated measure of ambulation, an activity that most people participate in on a daily basis, increasing the generalizability of our measure.

WOMAC pain was originally validated in a study of total knee and hip arthroplasty [10] so it makes sense that it performs well in those with late stage knee OA. This may also explain why WOMAC pain performed less well in our study in individuals with early disease. The WOMAC's lack of sensitivity to radiographic changes in early OA is a barrier to the development of novel disease modifying OA drugs (DMOADs) [6]. Given its ease of use and greater sensitivity to detect changes in ROA severity in individuals without established ROA at baseline, the AASK should be considered

for use in future research, particularly in clinical trials of early disease and in observational studies that include those with early disease. Trials of these treatments usually included only participants with established symptomatic knee OA, and once patients reach this point, they have already developed some increase in intra-articular stress [28]. Without modifying that factor, it would be difficult to change the natural history of the disease [28]. We postulate that the window of opportunity to modify the natural history of knee OA may occur prior to disease onset, or early in its development before mechanical derangement has occurred or when it can still be more readily modified. Similarly, we may find greater success in identifying disease modifying risk factors if we focus on those with earlier disease or without pre-existing disease, making validation of symptom instruments in early disease, such as AASK, a critical step in that process.

On the surface, it might seem that AASK and WOMAC function measure similar constructs. However, there is a subtle but important difference in these measures. WOMAC function assess the amount of difficulty people have doing particular activities while AASK assesses the amount of activity actually performed when considering symptom severity. Thus, AASK adds a dimension to a symptom outcome measure in knee OA that has not previously been considered.

We have previously published on PAKS (pain and activity knee symptom score) [14] a similar measure to the AASK, which assessed physical activity based on accelerometry data collected over a 7-day period. Consistent with the present study, we found that PAKS scores were more strongly associated with ROA severity than were WOMAC pain scores. An important benefit of AASK over PAKS is that the measure of physical activity is much simpler for AASK because PASE is simpler and less costly than use of accelerometry. The PASE

questionnaire can be completed in less than 10 min. In contrast, accelerometry data requires adherence to proper wearing of an expensive accelerometer over 7 days. Additionally, this must be processed extensively before it can be prepared in a usable form for analyses. Thus, from a logistic perspective, the AASK is easier to measure than PAKS. Unfortunately, we cannot comment on the relative sensitivity of PAKS versus AASK for detecting changes in ROA in people with knee OA as the OAI lacks sufficient longitudinal accelerometer data to address this question.

In summary, AASK has longitudinal validity as it relates to ROA severity. AASK is better able to discriminate a knee that transitions from a KL grade 0 to 1 than WOMAC. This is the finding we anticipated because AASK can discriminate two people who have the same WOMAC pain grade by adjusting for the estimated amount of ambulation participants engage in, thus scoring the more sedentary person as being more symptomatic. Because AASK performs well in both persons with and without established ROA and is simple to calculate, it should be considered when assessing symptoms of knee OA in observational studies and clinical trial that include those with early disease.

Funding information Dr. Lo is supported by K23 AR062127, an NIH/NIAMS funded mentored award; and Dr. Dunlop's awards, R01 AR054155 and P60 AR064464, provided support for design and conduct of the study, analysis, and interpretation of the data; and preparation and review of this work. This work is supported in part with resources at the VA HSR&D Center for Innovations in Quality, Effectiveness and Safety (No. CIN 13-413), at the Michael E. DeBakey VA Medical Center, Houston, TX. Dr. Hawker receives support as the Sir John and Lady Eaton Professor and Chair, Department of Medicine, University of Toronto. The Osteoarthritis Initiative is a public-private partnership comprised of five contracts (N01-AR-2-2258; N01-AR-2-2259; N01-AR-2-2260; N01-AR-2-2261; N01-AR-2-2262) funded by the National Institutes of Health, a branch of the Department of Health and Human Services, and conducted by the Osteoarthritis Initiative Study Investigators. Private funding partners include Merck Research Laboratories; Novartis Pharmaceuticals Corporation, GlaxoSmithKline; and Pfizer, Inc. Private sector funding for the Osteoarthritis Initiative is managed by the Foundation for the National Institutes of Health.

Compliance with ethical standards

Disclosures None.

References

- Cisternas MG, Murphy L, Sacks JJ, Solomon DH, Pasta DJ, Helmick CG (2016) Alternative methods for defining osteoarthritis and the impact on estimating prevalence in a US population-based survey. *Arthritis Care Res (Hoboken)* 68:574–580
- Nguyen US, Zhang Y, Zhu Y, Niu J, Zhang B, Felson DT (2011) Increasing prevalence of knee pain and symptomatic knee osteoarthritis: survey and cohort data. *Ann Intern Med* 155:725–732
- Kotlarz H, Gunnarsson CL, Fang H, Rizzo JA (2009) Insurer and out-of-pocket costs of osteoarthritis in the US: evidence from national survey data. *Arthritis Rheum* 60:3546–3553
- Mankin HJ (1993) Clinical features of osteoarthritis. In: Kelly WN, EDJ H, Ruddy S, Sledge CB (eds) *Textbook of Rheumatology*, 4th edn. W.B. Saunders Co., Philadelphia, PA, pp 1374–1384
- Guccione AA, Felson DT, Anderson JJ, Anthony JM, Zhang Y, Wilson PW, Kelly-Hayes M, Wolf PA, Kreger BE, Kannel WB (1994) The effects of specific medical conditions on the functional limitations of elders in the Framingham Study. *Am J Public Health* 84:351–358
- Lequesne M, Brandt K, Bellamy N, Moskowitz R, Menkes CJ, Pelletier JP et al (1994) Guidelines for testing slow acting drugs in osteoarthritis. *J Rheumatol Suppl* 41:65–71 discussion 72–63
- Neogi T, Zhang Y (2011) Osteoarthritis prevention. *Curr Opin Rheumatol* 23:185–191
- Bellamy N (1989) Pain assessment in osteoarthritis: experience with the WOMAC osteoarthritis index. *Semin Arthritis Rheum* 18:14–17
- Bellamy N (2002) WOMAC: a 20-year experiential review of a patient-centered self-reported health status questionnaire. *J Rheumatol* 29:2473–2476
- Bellamy N, Buchanan WW (1986) A preliminary evaluation of the dimensionality and clinical importance of pain and disability in osteoarthritis of the hip and knee. *Clin Rheumatol* 5:231–241
- Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW (1988) Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to anti-rheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 15:1833–1840
- Bellamy N, Wells G, Campbell J (1991) Relationship between severity and clinical importance of symptoms in osteoarthritis. *Clin Rheumatol* 10:138–143
- Altman R, Brandt KD, Hochberg MC, Moskowitz R (1996) Design and conduct of clinical trials in patients with osteoarthritis: recommendations from a task force of the Osteoarthritis Research Society. *Osteoarthr Cartil* 4:217–243
- Lo GH, McAlindon TE, Hawker GA, Driban JB, Price L, Song J et al (2015) Symptom assessment in knee osteoarthritis needs to account for physical activity level. *Arthritis Rheumatol* 67:2897–2904
- Jinks C, Jordan K, Croft P (2002) Measuring the population impact of knee pain and disability with the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC). *Pain* 100:55–64
- Dieppe PA, Cushnaghan J, Shepstone L (1997) The Bristol 'OA500' study: progression of osteoarthritis (OA) over 3 years and the relationship between clinical and radiographic changes at the knee joint. *Osteoarthr Cartil* 5:87–97
- Dieppe PA (2004) Relationship between symptoms and structural change in osteoarthritis. What are the important targets for osteoarthritis therapy? *J Rheumatol Suppl* 70:50–53
- Hannan MT, Felson DT, Pincus T (2000) Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. *J Rheumatol* 27:1513–1517
- Hawker GA, Stewart L, French MR, Cibere J, Jordan JM, March L, Suarez-Almazor M, Gooberman-Hill R (2008) Understanding the pain experience in hip and knee osteoarthritis—an OARSI/OMERACT initiative. *Osteoarthr Cartil* 16:415–422
- Davis AM, Perruccio AV, Canizares M, Hawker GA, Roos EM, Maillefert JF, Lohmander LS (2009) Comparative, validity and responsiveness of the HOOS-PS and KOOS-PS to the WOMAC physical function subscale in total joint replacement for osteoarthritis. *Osteoarthr Cartil* 17:843–847
- Washburn RA, Smith KW, Jette AM, Janney CA (1993) The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol* 46:153–162
- Peterfy C, Li J, Zaim S, Duryea J, Lynch J, Miaux Y, Yu W, Genant HK (2003) Comparison of fixed-flexion positioning with fluoroscopic semi-flexed positioning for quantifying radiographic joint-

- space width in the knee: test-retest reproducibility. *Skelet Radiol* 32: 128–132
23. Radiographic (x-ray) manual (version 2.1, 8-23-2006). <https://oai.epi-ucsf.org/datarelease/OperationsManuals.asp>. Accessed 17 Dec 2015
 24. Altman RD, Gold GE (2007) Atlas of individual radiographic features in osteoarthritis, revised. *Osteoarthritis Cartilage* 15(Suppl A): A1–A56
 25. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. *Biometrics* 33:159–174
 26. Project 15 test-retest reliability of semi-quantitative readings from knee radiographs. <https://oai.epi-ucsf.org/datarelease/ImageAssessments.asp>. Accessed 3 Feb 2017
 27. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD (1998) Knee Injury and Osteoarthritis Outcome Score (KOOS)—development of a self-administered outcome measure. *J Orthop Sports Phys Ther* 28:88–96
 28. Brandt KD, Dieppe P, Radin E (2009) Etiopathogenesis of osteoarthritis. *Med Clin North Am* 93:1–24 xv