Test-retest reliability, internal consistency, and discriminant validity of the Filipino version of Knee injury and Osteoarthritis Outcome Score among community-dwellers with knee osteoarthritis

Donald Manlapaz, Catherine Joy Escuadra, John Kenneth Ceazar Averia, Andrea Blancaflor, Rachel Ann Enriquez, Angela Mariz Ladeza, Angelica Marie Mandario, Jose Javier Mendoza, Thad Nuel Natividad

INTRODUCTION

Osteoarthritis (OA) is the most prevalent chronic joint disease and one of the most leading forms of pain and disability worldwide. Between the two types of lower extremity OA, knee OA continues to be more prevalent than hip OA across the globe and in the Asian region. In the Philippines, 11% of the population aged 60 and above have OA which is expected to double within the next 25 years. A cross-sectional study conducted in two different arthritis clinics in Metro Manila reported that out of 859 patients diagnosed with OA, 62.5% had knee OA while 1.6% had hip OA. With the evidence of elevated prevalence and burden of knee OA, focus on proper screening and assessment of the condition are essential in order to create an effective treatment and prevention management.

Keywords: KOOS, psychometric properties, knee osteoarthritis, outcome measures
The utilization of standardized Physical Therapy (PT) outcome measures is recommended to objectively evaluate the signs and symptoms of a given condition. Moreover, valid and reliable outcomes to measure a specific impairment, functional limitations, and quality of life (QoL) of the patient are critical at any stage of rehabilitation.

International groups and societies for OA recommend set of outcome measures to accurately classify and diagnose individuals with knee OA. Two of the most commonly used outcome measures are Knee injury and Osteoarthritis Outcome Score (KOOS) and Western Ontario and McMaster Universities (WOMAC). These outcome measures have been extensively used both in the clinical and research application.

The original KOOS was developed using WOMAC, which is intended to be used for knee injury that can subsequently result in posttraumatic OA. It is also widely used among individuals clinically and medically diagnosed with knee OA. It is composed of 42 questions divided into five separate subscales addressing knee symptoms, pain in activities of daily living (ADLs), function in sport and recreation, and knee-related QoL. It uses a five-point Likert scale scoring system ranging from 0 (least severe) to 4 (most severe) in order to answer each question.

KOOS has been extensively adapted in multiple languages and tested for validity and reliability, including a Filipino version making it a widely accepted outcome measure tool in assessing knee OA. The Filipino version of KOOS (F-KOOS) has well accepted cross-cultural adaptation and demonstrated acceptable psychometric properties such as Cronbach’s alpha, ICCs, item-to-domain correlations and validity in Filipino patients with knee OA. However, the psychometric properties such as test-retest reliability, internal consistency, and discriminant validity have not yet been investigated.

Interestingly, the F-KOOS has only been used among patients in hospitals and has yet been studied as a part of screening tools or outcome measures in the community setting (rural areas). This is important in the context of the community setting since most people who have knee OA do not seek medical care until the symptoms worsen due to financial constraints. In a developing country like the Philippines, various communities do not have access to immediate medical evaluation and treatment. The KOOS could be used as a convenient and accessible tool in assessing pain, ADLs, and function of individuals with OA of the knee.

Although the F-KOOS was already translated and culturally adapted, the identified psychometric properties such as test-retest reliability, internal consistency, and discriminant validity have not been explored. As per researchers’ knowledge, only one validity study for F-KOOS is available in the country. Despite the availability of the F-KOOS, there is a paucity in examining the applicability and usage of these tools in the community setting. Therefore, the primary aim of the study was to determine the test-retest reliability, internal consistency, and discriminant validity of the F-KOOS among the community-dwellers with knee OA. The secondary aim was to examine the suitability of the F-KOOS in terms of relevance and ease of understanding the tool.

**METHODS**

**Study Design**

A cross-sectional quantitative psychometric research design was utilized in this study. The participants were asked to attend two screening sessions, that were two weeks apart, in order to determine the test-retest reliability of F-KOOS. This study was reported in accordance with the Guidelines for Reporting Reliability and Agreement Studies (GRRAS) statement from the Enhancing the QUAlity and Transparency Of health Research (EQUATOR) network.

The study complied with the principles of the Declaration of Helsinki and Good Clinical Practice Guidelines of the Philippine Health Research Ethics Board. The ethical approval was obtained from the University of Santo Tomas-College of Rehabilitation Sciences Ethical Review Committee (Protocol Number: SI 2017-005).

**Sample size**

The sample size was based on COnsensus-based Standards for the selection of health Measurement Instrument (COSMIN) which is an appraisal tool used in evaluating the
methodological quality. According to the COSMIN tool, excellent (adequate sample size) is given greater than or equal to 100 recruited participants while poor for less than 30.

Participants

Community-dwellers aged 50 years old and above with knee pain or soreness were recruited for this study. Two licensed medical doctors were present during the assessment and data collection to identify participants with knee OA. The American College of Rheumatology (ACR) clinical criteria using history and physical examination for classification of knee OA was used as the basis for the inclusion of participants. Participants who understood Filipino/Tagalog written and verbal instructions were recruited since the F-KOOS is a self-administered outcome measure. Participants with health conditions that affected the level of independence in ADLs were excluded from the study. Since the tool assesses pain, function, and QOL, the study excluded any conditions (e.g. Acute neurologic conditions, fracture, sprain, etc.) that influence these factors other than knee OA.

Recruitment and study setting

Participants were recruited from Binangonan, Rizal through community advertising. Prior to advertising, several consultations were made with the Community-Based Rehabilitation (CBR) head, Barangay Officials and, Health City Administrators. The permission and assistance were sought from the City Mayor’s office through the CBR head. The study was supported by government health officials of Binangonan, Rizal and commenced upon receiving an approval.

Communication to respective barangay officials was done to identify the time and location of the data gathering.

Recruitment took place from November 2017 to December 2017. Potential participants from the communities were invited to the city’s Municipal Hospital and Convention Hall where they were assessed.

Instrument

The F-KOOS is a 42-item self-administered outcome measure consisting of five subscales: Pagkirot (9 items), Sintomas (7 items), Pang-araw-araw na gawain (17 items), Gampanin, Isports at Libangan (5 items), and Kalidad ng buhay (4 items). After rating each item, the scores of each subscale are then individually converted to a 0-100 scale (0 = extreme knee problems, 100 = no knee problems) wherein the lower score would mean a more severe condition of knee OA. The F-KOOS was translated by a qualified instructor from the University of the Philippines, and another independent translator who is knowledgeable of the KOOS.

To determine if the questions of the F-KOOS are suitable and relevant to the condition of the participants, a survey was distributed after the administration of the F-KOOS. According to a study by Caudle and colleagues, the satisfaction can be determined by the following domains: appropriateness, convenience (easy to apply), and comprehension (perception). This survey was done in the form of a Likert scale. There is no gold standard as to how to assess the satisfaction of instrument or an outcome measure, but previous studies have focused on its relevance to evidence-based practice in terms of developing better health outcomes.

Procedure

During the data collection, the researcher explained the purpose and procedures of the study to the participant. The assessor administered the consent process and answered any queries the participants had. An assessment tool kit, which contained the F-KOOS and a self-administered participants’ satisfaction questionnaire was prepared and given to participants.

The participants answered the questionnaire for approximately 10-15 minutes. The participants were instructed to drop the accomplished questionnaires in a secured ballot box. The ballot box was accessed and kept by one (1) researcher until the period of data analysis. After two weeks, the F-KOOS was re-administered by the researcher to assess the test re-test reliability.

Data Analysis

Stata 15 (Serial Number: 401506343769) was used for data analyses. Descriptive statistics (i.e. mean and standard deviation) were used to
determine baseline characteristics of participants with and without knee OA and F-KOOS subscale scores. Frequencies and percentages were used to determine the preferred questionnaire by the participants, the answers were reported through descriptive statics. To evaluate the variability of the responses of the participant’s floor and ceiling effects were computed. The test-retest reliability and discriminant validity were both analyzed through inferential statistics. The test-retest reliability was tested using the Intraclass correlation Coefficient (ICC). It reflects both systematic and random differences in the test scores of the first and second questionnaire administered and thus, values of ICC may vary from 0 (unreliable) to 1 (perfectly reliable). The ICC was chosen in preference to the Pearson correlation, which may overestimate reliability. The standard error of measurement (S.E.M.) and minimal detectable change (MDC) was computed. Discriminant validity was assessed by comparing the results of participants with knee pain diagnoses with knee OA versus participants without knee OA. Between-group analyses were done by using independent t-test per F-KOOS subscale. P-values less than 0.05 were considered significant in the analyses.

RESULTS

Participants

A total of 53 participants with the age of 69.67 ± 5.83 years old (35 female, 18 males) were included in the study. 41 participants were diagnosed by medical doctors to have knee OA according to ACR clinical criteria without other comorbidities. Figure 1 shows the F-KOOS subscale scores of the participants for the 1st assessment (baseline) and 2nd assessment (re-test). Table 1 shows the results of F-KOOS scores showing the mean, median, range scores of each F-KOOS subscale and their corresponding floor and ceiling effect.

Figure 1: Results of F-KOOS of participants with knee OA at baseline and re-test (n=41)

Note: ADL- Activity of daily living; QoL- Quality of life
Table 1: Summary of F-KOOS responses of participants with knee OA at baseline assessment (n=41)

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Items</th>
<th>Mean (SD)</th>
<th>Median</th>
<th>Range</th>
<th>Floor (n,%)</th>
<th>Ceiling (n,%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain</td>
<td>9</td>
<td>59.71 (22.13)</td>
<td>63.89</td>
<td>25.00-97.22</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Symptoms</td>
<td>7</td>
<td>57.19 (21.32)</td>
<td>53.57</td>
<td>17.86-96.43</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ADL</td>
<td>17</td>
<td>61.39 (20.85)</td>
<td>63.24</td>
<td>29.41-95.59</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sports/Recreation</td>
<td>5</td>
<td>51.19 (26.45)</td>
<td>55.00</td>
<td>0-95.00</td>
<td>2 (4.88)</td>
<td></td>
</tr>
<tr>
<td>QOL</td>
<td>4</td>
<td>50.46 (23.28)</td>
<td>43.75</td>
<td>18.75-100.0</td>
<td>-</td>
<td>1 (2.44)</td>
</tr>
</tbody>
</table>

Note: SD- standard deviation; n- count; % - percentage; ADL- Activity of daily living; QoL- Quality of life

Test-retest reliability

Table 2 shows the result of the test-retest analysis. With the alpha set at 0.05, p-values of the domains of the F-KOOS in the pre-test and re-

Table 2: Comparison of F-KOOS responses across time of participants with knee OA (n=41)

<table>
<thead>
<tr>
<th>F-KOOS subscale</th>
<th>Items</th>
<th>Pre-test Mean (SD)</th>
<th>Re-test Mean (SD)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain</td>
<td>9</td>
<td>57.02 (3.19)</td>
<td>55.77 (3.15)</td>
<td>0.78</td>
</tr>
<tr>
<td>Symptoms</td>
<td>7</td>
<td>59.82 (3.30)</td>
<td>54.91 (3.37)</td>
<td>0.30</td>
</tr>
<tr>
<td>ADL</td>
<td>17</td>
<td>61.88 (3.12)</td>
<td>58.46 (3.00)</td>
<td>0.31</td>
</tr>
<tr>
<td>Sports/Recreation</td>
<td>5</td>
<td>51.83 (3.98)</td>
<td>49.49 (4.17)</td>
<td>0.69</td>
</tr>
<tr>
<td>QOL</td>
<td>4</td>
<td>50.58 (3.50)</td>
<td>48.03 (3.33)</td>
<td>0.60</td>
</tr>
</tbody>
</table>

Note: SD- standard deviation; n- count; ADL- Activity of daily living; QoL- Quality of life

Table 3 which shows the ICC, 95% CI, standard error of the mean, minimal detectable change, and Cronbach alpha for F-KOOS subscales. The ICC showed slightly good level of agreement from 0.64-69.

Table 3: Test-retest reliability and internal consistency of F-KOOS among participants with knee OA (n=41)

<table>
<thead>
<tr>
<th>Subscales</th>
<th>Items</th>
<th>Cronbach's alpha</th>
<th>ICC (95% CI)</th>
<th>SEM</th>
<th>MDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain</td>
<td>9</td>
<td>0.92</td>
<td>0.69 (0.40-0.84)</td>
<td>2.36</td>
<td>6.52</td>
</tr>
<tr>
<td>Symptoms</td>
<td>7</td>
<td>0.84</td>
<td>0.69 (0.40-0.84)</td>
<td>2.23</td>
<td>6.16</td>
</tr>
<tr>
<td>ADL</td>
<td>17</td>
<td>0.96</td>
<td>0.69 (0.40-0.84)</td>
<td>2.17</td>
<td>6.00</td>
</tr>
<tr>
<td>Sports/Rec</td>
<td>5</td>
<td>0.89</td>
<td>0.64 (0.30-0.81)</td>
<td>2.86</td>
<td>7.90</td>
</tr>
<tr>
<td>QOL</td>
<td>4</td>
<td>0.87</td>
<td>0.64 (0.30-0.82)</td>
<td>2.42</td>
<td>6.69</td>
</tr>
</tbody>
</table>

Note: ICC- intraclass correlation; CI- confidence interval; SEM- standard error of mean; MDC- minimal detectable change; ADL- Activity of daily living; QoL- Quality of life

Internal consistency

Table 3 shows the result of internal consistency. All domains of the F-KOOS have a Cronbach alpha >0.70 which indicates high internal consistency within all domains ranging from 0.87 to 0.96.

Discriminant validity

To determine discriminant validity, Table 4 reports differences in F-KOOS subscale scores between participants with and without knee OA.
Table 4: Discriminant validity of F-KOOS among participants with and without knee OA

<table>
<thead>
<tr>
<th></th>
<th>With knee osteoarthritis (n=41)</th>
<th>Without knee osteoarthritis (n=12)</th>
<th>t value</th>
<th>p-value</th>
<th>d-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain</td>
<td>85.42 ± 18.12</td>
<td>59.71 ± 22.13</td>
<td>3.67</td>
<td>&lt;0.01*</td>
<td>1.21</td>
</tr>
<tr>
<td>Symptoms</td>
<td>77.68 ± 17.50</td>
<td>57.19 ± 21.32</td>
<td>3.04</td>
<td>&lt;0.01*</td>
<td>1.00</td>
</tr>
<tr>
<td>ADL</td>
<td>83.59 ± 19.48</td>
<td>61.39 ± 20.85</td>
<td>3.29</td>
<td>0.02*</td>
<td>1.08</td>
</tr>
<tr>
<td>Sports/Rec</td>
<td>80.42 ± 24.26</td>
<td>51.19 ± 26.45</td>
<td>3.43</td>
<td>&lt;0.01*</td>
<td>1.12</td>
</tr>
<tr>
<td>QOL</td>
<td>83.33 ± 23.59</td>
<td>50.46 ± 23.28</td>
<td>4.29</td>
<td>&lt;0.01*</td>
<td>1.41</td>
</tr>
</tbody>
</table>

Note: *significant at p<0.05; SD- standard deviation

Suitability of F-KOOS

In terms of suitability, only 40 participants answered the survey. Most participants who answered the suitability questionnaire agreed (strongly agreed 65% and agreed 30%) that the outcome measure was relevant. Five percent disagreed, 55% agreed, and 45% of the participants strongly agreed that the outcome measure was appropriate. Forty-five percent of the participants agreed and 55% of participants strongly agreed that the outcome measure was easy to understand and answer. Figure 2 summarizes the response rate regarding the suitability of F-KOOS.

Figure 2: Summary of perceptions of participants regarding use of F-KOOS (n=40)

DISCUSSION

This study demonstrated that the F-KOOS has acceptable test-retest reliability, good internal consistency, and discriminant validity in individuals with knee OA. Suitability of F-KOOS for community-dwelling patients showed that majority of the participants agreed that the F-KOOS was easy to answer and understand. Participants agreed that the F-KOOS was relevant to their experience and condition.

Psychometric properties of F-KOOS

Test-retest reliability refers to the extent to which test results are consistent over time. Findings in this study have established that the F-KOOS has an acceptable test-retest reliability coefficient for all subscales in the present study.
particularly in pain and ADL with an ICCs ranging from 0.64 to 0.69. This suggests satisfactory stability of F-KOOS scores over time making the version of the tool reliable for obtaining results among knee OA community dwellers. The result is also comparable with the findings in studies done in other languages with similar conditions including the original KOOS with ICC of 0.75 to 0.93.9

The Internal consistency, described by computing Cronbach alpha of F-KOOS, was acceptable in all domains (pain, other symptoms, function in sports and recreation, function in daily living, and knee-related QoL) ranging from 0.87 to 0.96, which exceeded the normative value of 0.70, and reflective of the original version of KOOS9. This indicated a high correlation of items in contrast to Persian KOOS with a low Cronbach Alpha of 0.25 in Symptoms domain and 0.65 in QoL domain.21 Singapore Chinese version also has a below-average results in Symptom domain which is 0.65 and Pain domain which is 0.64.15

The disparity in results may be attributed to several factors such as characteristics of participants, and logistics in the administration of the tool Korean KOOS, on the other hand, presents a good Cronbach alpha in all domains ranging from 0.7 to 0.9522 as well as the Italian KOOS with 0.7 to 0.95.23

The study also included assessment discriminant validity among the other variables to be evaluated and by obtaining the p values (p<0.05), it can be determined that there is a significant difference between the participants with and without knee OA. Thus, the F-KOOS can effectively distinguish between those adults with knee OA and without knee OA and can be used in the Filipino community setting.

Suitability of F-KOOS in the community-dwelling knee OA

Majority of the participants (95%) of our study either agreed or strongly agreed that the KOOS is indeed relevant and appropriate to their present condition. All participants either agreed or strongly agreed that the KOOS was easy to understand and answer. These results corroborate with a study by Caudle and colleagues where the preference of outcome measures was established.16 In this current study, the F-KOOS was well-received as an outcome measure that can be applicable to their condition. The participants also did not have any difficulty in answering the F-KOOS and understanding its contents in the Filipino community setting.

Psychometric tests can be administered to a large group of people at a time, without having to tailor each one to different individuals. It should be noted that this increases the speed and ease of administration of the outcome measure; although, much of the value of any test depends on the administrators. Tests can be poorly presented or explained, which can cause the results to not be accurate. Moreover, because the emphasis is often placed on the results of these psychometric tests, it can be potentially damaging to the study, especially to those who hail from different language and cultural backgrounds.

Limitations of the study

The findings of this study have several methodological limitations. Although our sample size is moderate to high, the sample was composed of those with knee pain and with knee OA diagnosis. Aside from this, the study utilized self-reported data only making results possibly prone to certain biases. Further studies are needed to clearly ascertain differences between the group with and without knee OA using other external or criterion-related validity of the objective tests and clinical tests.

The group encountered some limitations during the course of the study: scheduling of the data gathering. Nevertheless, the researchers were able to address these problems and gather & analyze the data needed. In terms of scheduling, those who did not come back during the second phase were asked to answer in a more convenient time; their respective assessment tools were given to the barangay coordinator and was administered by local health workers. During the administration of the test, the group encountered some of the participants that had difficulty in reading the outcome measure; the researchers opted to help them by reading the questions and choices to them for better understanding and appreciation of the outcome measure tool.
Implication to practice and research

The F-KOOS can be used easily as an assessment tool in the community setting as it is convenient to administer and would not require additional materials and training. Outcome questionnaires are important in assessment to objectively evaluate the signs, symptoms, and function of patients. It is also used to evaluate the progress of patients during the course of treatment. Through the results of this study, PT can be confident to use the F-KOOS in screening individuals with knee OA in the community setting.

The KOOS has been extensively adapted into other languages with its psychometric properties evaluated. There is only one existing study done in the Philippines regarding the F-KOOS. With this study, the F-KOOS has now been tested to have acceptable test-retest reliability, good internal consistency, and discriminant validity. The suitability of the F-KOOS demonstrated very high in the majority of the participants who answered the survey. Thus, contributing to the body of literature in assessing individuals with knee OA using F-KOOS. Future researches about the tool may focus on other psychometric properties of the tool such as criterion validity (when compared with clinical tools), construct validity, sensitivity, and specificity.

CONCLUSION

This study demonstrated that the F-KOOS has acceptable test-retest reliability, good internal consistency, and discriminant validity in individuals with knee OA. The study further determined that the use of the F-KOOS is appropriate, relevant, and easy to understand. Health care providers including PT can be assured that they are evaluating individuals with knee OA using a valid and reliable tool which can lead to creating an effective treatment and prevention strategies.

Conflict of Interest

The authors declare no conflict of interest.

References


