Comparative, Validity and Responsiveness of the HOOS-PS and KOOS-PS to the WOMAC Physical Function Subscale in Total Joint Replacement for Osteoarthritis

Aileen M. Davis¹, Anthony V. Perruccio²,
Mayilee Canizares³, Gillian A. Hawker⁴, Ewa M. Roos⁵,
Jean-Francis Maillefert⁶ and L. Stefan Lohmander⁷

1. Division of Health Care and Outcomes Research and Arthritis Community Research and Evaluation Unit, Toronto Western Research Institute; Departments of Physical Therapy, Rehabilitation Science, Health Policy, Management and Evaluation and Institute of Medical Science, University of Toronto, Toronto, Canada
2. Division of Health Care and Outcomes Research and Arthritis Community Research and Evaluation Unit, Toronto Western Research Institute; Dalla Lana School of Public Health, University of Toronto, Toronto, Canada
3. Division of Health Care and Outcomes Research and Arthritis Community Research and Evaluation Unit, Toronto Western Research Institute, Toronto, Canada
4. Division of Rheumatology, Department of Medicine, Women’s College Hospital, Departments of Health Policy, Management and Evaluation and Department of Medicine, University of Toronto, Toronto, Canada
5. Institute of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark and Department of Orthopaedics, Clinical Sciences Lund, Lund University, Sweden
6. Dijon University Hospital and INSERM U887, University of Burgundy, Dijon, F-21079, France
7. Department of Orthopaedics, Clinical Sciences Lund, Lund University, Lund, Sweden

**Corresponding Author:**
Aileen M. Davis, PhD
Senior Scientist
Division of Health Care and Outcomes Research and Arthritis Community Research and Evaluation Unit
Toronto Western Research Institute
399 Bathurst Street - MP11-322
Toronto, Ontario M5T 2S8
Tel: 416 603-5543
Fax: 416 603-6288
Email: adavis@uhnresearch.ca
Objective: To evaluate the internal consistency of the Hip disability and Osteoarthritis Outcome Score- Physical Short (HOOS-PS) and the Knee injury and Osteoarthritis Outcome Score-Physical Short (KOOS-PS) in total hip (THR) and total knee (TKR) replacement. Construct validity and responsiveness were compared to the WOMAC Likert 3.0 physical function subscale (PF) and the PF excluding the items in the short measures (PF-exclusions).

Methods: Participants completed the full HOOS or KOOS, measures of fatigue, anxiety, depression and the Chronic Pain Grade (CPG) pre surgery and the HOOS or KOOS 6 months post surgery. Internal consistency for the HOOS-PS and KOOS-PS was calculated using Cronbach’s alpha. For construct validity, it was hypothesized that correlations between the HOOS-PS or KOOS-PS and PF and PF-exclusions with fatigue, CPG, anxiety and depression and HOOS/KOOS pain scales would differ by magnitudes of < 0.1. Standardized response means (SRM) were calculated for the HOOS-PS, KOOS-PS, PF and PF-exclusions and hypothesized to be >1.

Results: The THR group (n=201) had a mean age of 62.3 years; 53.2% were female. The TKR group (n=248) had a mean age of 64.5 years; 63.7% were female. Cronbach’s alpha was 0.79 and 0.89 for the HOOS-PS and KOOS-PS respectively confirming that the measures represented a homogeneous construct. The correlation of the HOOS-PS to the PF and PF-exclusions was 0.90 and 0.86 respectively; r=0.90 (PF) and r=0.85 (PF-exclusions) for the KOOS-PS. The results supported the construct validity hypotheses. For THR, the SRM was 1.5, 1.7 and 1.7 for the HOOS-PS, PF and PF-exclusions; for TKR, the SRM was 1.4, 1.5 and 1.7, respectively.

Conclusions: The short HOOS-PS and KOOS-PS represent homogenous short measures of
physical function with similar construct validity and responsiveness to the 17-item PF. The HOOS-PS and KOOS-PS are parsimonious, valid and responsive for evaluating physical function in THR and TKR.
Introduction (Word Count 2346)

Osteoarthritis (OA) is one of the most prevalent diseases in the developed world and is a major cause of pain and physical disability\cite{1,2}. It is most common in the hip and knee, is a leading cause of activity limitation, loss of independence, decreased quality of life and is a significant economic burden in terms of health care costs\cite{1,3-6}. Hence, change in physical function as a result of intervention is a critical outcome. The Western Ontario McMaster Universities’ Osteoarthritis Index (WOMAC)\cite{7,8} is one of the most commonly used outcome measures; it includes a physical function subscale (PF) of demonstrated reliability, validity and responsiveness for people with hip and or knee OA. However, the subscale has 17 items and, when clinicians or researchers are using a battery of measures they often require parsimonious measures to limit respondent burden. Additionally, concerns have been expressed that the WOMAC physical function subscale assesses a limited range of functional disability and has redundant items\cite{9-11}. Given these potential limitations of the WOMAC PF, members of our group developed short measures of physical function for hip and knee osteoarthritis\cite{12,13}.

The HOOS- Physical Function Short-form (HOOS-PS) and KOOS-Physical Function Short-form (KOOS-PS) were developed from the Activities of Daily Living subscale (which subsumes the 17 physical function items of the WOMAC Likert 3.0) and the Sport and Recreation subscales of the Hip disability and Osteoarthritis Outcome Score (HOOS) and the Knee injury and Osteoarthritis Outcome Score (KOOS), respectively\cite{14-16}. The items included in the HOOS-PS and KOOS-PS are presented in Table 1. Having been developed by ensuring fit of the data to the Rasch model, these short measures provide interval (as opposed to ordinal) level scores that can appropriately be subjected to inferential statistics\cite{12,13}. Additionally, the measures represent a measure solely of physical function by virtue of their meeting the
requirements of strict unidimensionality\textsuperscript{[12,13]}. However, these short measures were purposely developed to represent the spectrum of hip and knee OA and included those from community samples, those who received conservative management, surgery other than total joint replacement (e.g. osteotomy), as well as those who were scheduled for hip or knee replacement surgery. The measures have not yet been tested for responsiveness. Given that total hip (THR) and total knee replacement (TKR) are the treatment of choice with known effectiveness for people with end stage arthritis, the measurement properties of these short measures need to be evaluated in this patient group.

The purpose of this study was to evaluate the internal consistency of the HOOS-PS and KOOS-PS and to evaluate their construct validity and responsiveness as compared to the WOMAC Likert 3.0 physical function subscale in people undergoing THR and TKR.

**Methods**

This study included 201 people who had primary THR and 226 people who had primary TKR for OA who had 6 months of follow-up post surgery. The surgery occurred at one of four academic hospitals in Toronto, Canada. Individuals were eligible for inclusion in the study if they were over the age of 18 years, were undergoing primary THR or TKR for OA, were able to read and comprehend English in order to complete the questionnaires and consented to participate. Exclusion criteria included joint replacement for inflammatory arthritis, fracture, tumour or acute trauma, hemi-arthroplasty and revision arthroplasty. The study was approved by the human subject review board of each of the participating institutions.

The participants completed a battery of self-report questionnaires by mail within two weeks prior to their joint replacement surgery; the full HOOS or KOOS, a measure of fatigue
from the Profile of Mood States (POMS) \cite{17}, the Hospital Anxiety and Depression Scale (HADS) \cite{18} and the Chronic Pain Grade (CPG) \cite{19-21}. The HOOS or KOOS also was completed at 6 months post surgery.

The HOOS-PS and KOOS-PS scores were derived from the responses to full HOOS and KOOS as accrual to the sample began in April 2006, prior to the development of the short measures. The HOOS-PS consists of 5 items and the KOOS-PS has 7 items; both are scored zero to 100 with zero scores representing no difficulty \cite{12,13}. Similarly, the WOMAC Likert 3.0 17 item PF subscale was extracted from the HOOS/KOOS respectively and summed to create a 0 to 68 score in which zero represented no difficulty. Given that the HOOS-PS and KOOS-PS include items from the WOMAC, we also created a physical function score based on only the items of the WOMAC that were not included in the short measures (PF-exclusions). The PF-exclusion score ranged from 0 to 56 for the THA group and 0 to 48 for the TKA group with zero representing no difficulty. The PF-exclusion scores avoid over-estimation of the correlations calculated for construct validity and of the standardized response mean for responsiveness in comparison to the short measures.

The WOMAC Likert 3.0 pain scale, the POMS fatigue subscale, the CPG and the anxiety and depression scores from the HADS were used for testing construct validity. The WOMAC Likert 3.0 pain scale \cite{7} was extracted from the HOOS or KOOS. The score ranges from 0 to 20. The Profile of Mood States (POMS) fatigue subscale includes 5 items with a total score ranging from 0 to 20 \cite{17}. Both the anxiety and depression subscale scores range from 0 to 21 \cite{18}. The CPG score ranges from 0 to 27 \cite{19-21}. For all of these measures zero represents no symptoms. These measures all have reported reliability and validity \cite{7,17-22,2}. 


Descriptive statistics were calculated for sample characteristics as appropriate to the type of data. Cronbach’s alpha was calculated as a measure of internal consistency \[^{23}\]. As a test of construct validity, it was hypothesized that correlations between each of the measures of physical function (HOOS-PS or KOOS-PS, PF and PF-exclusions) and the POMS, CPG, HADS anxiety and depression and WOMAC pain scales would differ by magnitudes of less than 0.1. Additionally, we expected the correlations among the measures of physical function (HOOS-PS or KOOS-PS, PF and PF-exclusions) to be higher than with the correlations with the pain, fatigue and anxiety and depressions measures. Based on the distribution of the data, Pearson correlation coefficients were calculated. The standardized response mean (SRM) \[^{24}\] was calculated for each of the HOOS-PS or KOOS-PS and for the PF and PF-exclusions and was hypothesized for all measures to be large and greater than 1. Analyses were conducted separately for the THR and TKR participants.

**Results**

The mean age of the THR group was 62.3 years (range 31-86) whereas the mean age of the TKR group was two years older at 64.5 years. Just over 53% were female in the THR group compared with 63% in the TKR group. Table 2 presents the sample characteristics. These data are similar to the description of those individuals reported in the Canadian Joint Replacement Registry for those surgeons reporting to the Registry \[^{25}\].

As assessed by Cronbach’s alpha, the internal consistency of the HOOS-PS was 0.79 and 0.89 for the KOOS-PS confirming that the measures represented a homogeneous construct. Table 3 shows the correlation coefficients for the tests of construct validity and demonstrates that the findings for the HOOS-PS and KOOS-PS are similar. For all the measures of physical function,
the correlations are highest with the WOMAC pain subscale, ranging from 0.70 to 0.80. In contrast, the correlations for all physical function measures with fatigue, CPG and depression subscales are moderate ranging from 0.33 to 0.66; for anxiety the correlations for the HOOS-PS was 0.19 as compared to 0.38 for the KOOS-PS. In keeping with the hypotheses, the physical function measures correlated with the given constructs (i.e. each of WOMAC pain, fatigue, the CPG, anxiety and depression) within 0.10. As hypothesized, the associations among the physical function measures were higher than among the pain, fatigue and anxiety and depression measures (i.e. associations for similar constructs were higher than for dissimilar constructs). The correlations of the HOOS-PS to the PF and PF-exclusions were 0.90 and 0.86, respectively. The KOOS-PS was highly correlated with the PF (r=0.90) and the PF-exclusions (r=0.85).

From pre surgery to 6 months post surgery, both the THR and TKR groups had significant and large improvements in physical function as measured by the short measures, the WOMAC PF and the PF-exclusions (Table 4). For the THR group, the SRM ranged from 1.5 to 1.7 and for the KOOS-PS from 1.4 to 1.7. Again our a priori hypothesis was supported by these data.

**Discussion**

This work has demonstrated that the 5 item HOOS-PS and 7 item KOOS-PS have similar construct validity and responsiveness to the WOMAC Likert 3.0 17 item physical function subscale within a sample of people undergoing total hip or total knee replacement. In addition to limiting response burden, particularly when a battery of measures evaluating different constructs is used, the HOOS-PS and KOOS-PS have the advantage of including more demanding activities such that they can be used across the spectrum of severity of hip and knee OA.
Prior work has suggested that the WOMAC physical function subscale includes redundant items that provide little additional information [9-11]. This was supported in the current study as the correlations among the three measures of physical function were high (range 0.85 to 0.90), irrespective of the hip or knee sample, suggesting that the measures are providing similar information. Most notably, the correlation of the HOOS-PS and KOOS-PS to PF and PF-exclusions differed by only 0.04 (hip) and 0.05 (knee).

Given the high correlation of the physical measures, the relationships of the HOOS-PS and KOOS-PS to the constructs of WOMAC pain, chronic pain, fatigue, anxiety and depression should be similar in magnitude to those reported in the literature for the 17 item WOMAC physical function subscale. Similar to this study, the data for hip and knee OA in general supports that the WOMAC physical function subscale is moderately correlated with measures of pain, fatigue, and mood in community samples and in patients managed by non-surgical modalities [7,22,26,27]. However, it should be noted that these studies, while using measures of similar constructs, did not use the same measures. There is little specific data for total joint replacement, particularly at the evaluation times of the current study, but the data similarly report moderate correlations ranging from 0.4 to 0.6 [28-30]. It should be noted that the correlation with the WOMAC Likert 3.0 pain subscale is higher (magnitudes of 0.73 and above) than the correlation with the Chronic Pain Grade in this study. Terwee et al. similarly reported an association of 0.74 between the WOMAC pain and physical function subscales in people with TKR [30]. This higher correlation is not surprising as the association of the WOMAC pain subscale and the physical function subscale is likely confounded by asking about pain on specified activities [31].
The large SRMs for the HOOS-PS and KOOS-PS indicate the effectiveness of total hip and knee replacement further suggest that these short measures maintain their psychometric properties. For the THR group, the SRM differed among the measures by 0.2 with the HOOS-PS having the smallest SRM of 1.5 compared to the PF and PF-exclusions; the KOOS-PS SRM differed by 0.3 with a magnitude of 1.4. These smaller SRMs for the HOOS-PS and KOOS-PS likely result from the more difficult items demonstrating less change in the joint replacement group. For example, people undergoing joint replacement may not ‘run’, an item in the HOOS-PS, such that this item in the questionnaire may not reflect change following surgery or change in only a portion of the group. However, emerging research suggests that people are looking to be able to participate in more than routine activities of daily living following joint replacement \[32\]. Additionally, there is controversy over what recreational activities are considered safe following hip or knee replacement \[33\]. Hence, the relevance of including these higher demand activities in outcome measures is yet to be determined.

While the studies in the literature similarly report large responsiveness statistics, comparison to this study is limited by the variability in the time at which responsiveness is reported and by the variability in how responsiveness is calculated. Wright et al. have shown that different methods of calculating responsiveness result in different magnitudes and interpretation \[33\]. In this study, we reported responsiveness at six months post surgery and used the standardized response mean, thereby accounting for the paired nature of the data. The studies by Wright et al. and Lingard et al. in patients with primary THR and TKR, respectively, both reported a SRM of 1.4 at one-year post surgery \[34,35\]. It should be noted that the study by Wright et al. included those undergoing primary THR for other than OA. In research, the critical issue of the magnitude of the SRM is related to sample size requirements as the larger the effect, the
fewer study participants who are required for a given level of power. However, given how large the SRM is for the HOOS-PS and KOOS-PS respectively as compared to the WOMAC physical function subscale, the effect on sample size is negligible.

The major limitation of this study is that, since the short measures did not exist at the time of the inception of the cohort, the HOOS-PS and KOOS-PS were extracted from the full HOOS and KOOS as opposed to their completion in random order as measures separate from the full measure. Additionally, through a mailed survey it is not possible to prevent individuals from looking to prior answers given that the short measures contain the items with identical wording even had the short and longer versions of the questionnaires been used. Interviewer administration would be required.

In summary, this first study to our knowledge of the HOOS-PS and KOOS-PS in people with total hip or knee replacement provides evidence of construct validity and responsiveness of the measures as compared to the longer WOMAC Likert 3.0 physical function subscale. While further testing in additional samples is required, the evidence from this work suggests that these short measures are viable and maintain their psychometric quality for use in joint replacement, particularly when respondent burden and feasibility are of concern.
Acknowledgments

This work was supported in part by an operating grant from the Canadian Institutes of Health Research, grant number 77518. The funding agency had no role in the conceptualization or execution of the project nor did the agency contribute to the analysis or interpretation of the results.
References


30. Terwee CB, van der Slikke RMA, van Lummel RC, Benink RJ, Meijers WJH, and de Vet HCW. Self-reported physical functioning was more influenced by pain than performance-based physical functioning in knee-osteoarthritis patients. Journal of Clinical Epidemiology 2006; 59 (7):724-731.


Table 1: Items in the HOOS-PS and KOOS-PS and their subscale of origin

<table>
<thead>
<tr>
<th>Subscale of Origin from HOOS or KOOS</th>
<th>HOOS-PS (5 items)</th>
<th>KOOS-PS (7 items)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activities of Daily Living subscale*</td>
<td>Sitting</td>
<td>Rising from bed</td>
</tr>
<tr>
<td></td>
<td>Descending stairs</td>
<td>Putting on sock/stockings</td>
</tr>
<tr>
<td></td>
<td>Getting in/out of bath/shower</td>
<td>Rising from sitting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bending to the floor</td>
</tr>
<tr>
<td>Sport and Recreation subscale</td>
<td>Running</td>
<td>Twisting/pivoting on your injured knee</td>
</tr>
<tr>
<td></td>
<td>Twisting or pivoting on your loaded leg</td>
<td>Kneeling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Squatting</td>
</tr>
</tbody>
</table>

HOOS= Hip disability and Osteoarthritis Outcome Score- Physical Short
KOOS= Knee injury and Osteoarthritis Outcome Score-Physical Short
*subscale subsumes the 17 physical function items of the WOMAC Likert 3.0
Table 2: Characteristics of the study sample undergoing primary total hip or total knee replacement for OA

<table>
<thead>
<tr>
<th></th>
<th>Total Hip Replacement (n=201)</th>
<th>Total Knee Replacement (n=248)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age: mean (sd)</td>
<td>62.3 (12.1)</td>
<td>64.5 (10.3)</td>
</tr>
<tr>
<td>Sex: M:F</td>
<td>1:1.3</td>
<td>1:1.7</td>
</tr>
<tr>
<td>BMI: &lt;25</td>
<td>4 (2%)</td>
<td>1 (0.4%)</td>
</tr>
<tr>
<td>25-29</td>
<td>61 (30.3%)</td>
<td>36 (14.5%)</td>
</tr>
<tr>
<td>30-34</td>
<td>66 (32.8%)</td>
<td>70 (28.2%)</td>
</tr>
<tr>
<td>≥35</td>
<td>61 (4.5%)</td>
<td>101 (40.7%)</td>
</tr>
<tr>
<td>missing</td>
<td>9 (4.5%)</td>
<td>40 (16.1%)</td>
</tr>
<tr>
<td>Marital status: single</td>
<td>18 (9.0%)</td>
<td>22 (8.9%)</td>
</tr>
<tr>
<td>Married or living with someone</td>
<td>183 (91.0%)</td>
<td>204 (82.2%)</td>
</tr>
<tr>
<td>missing</td>
<td>0</td>
<td>22 (8.9%)</td>
</tr>
</tbody>
</table>

BMI=body mass index in kg/m²
Table 3: Construct Validity of the HOOS-PS and KOOS-PS: Pearson’s Correlation Coefficients

<table>
<thead>
<tr>
<th></th>
<th>WOMAC Pain</th>
<th>Chronic Pain Grade</th>
<th>POMS Fatigue</th>
<th>HADS Anxiety</th>
<th>HADS Depression</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Hip Replacement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOOS-PS</td>
<td>0.70</td>
<td>0.56</td>
<td>0.38</td>
<td>0.19</td>
<td>0.36</td>
</tr>
<tr>
<td>PF</td>
<td>0.80</td>
<td>0.62</td>
<td>0.40</td>
<td>0.19</td>
<td>0.35</td>
</tr>
<tr>
<td>PF-exclusions</td>
<td>0.80</td>
<td>0.62</td>
<td>0.38</td>
<td>0.19</td>
<td>0.33</td>
</tr>
<tr>
<td><strong>Total Knee Replacement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOOS-PS</td>
<td>0.73</td>
<td>0.56</td>
<td>0.42</td>
<td>0.39</td>
<td>0.42</td>
</tr>
<tr>
<td>PF</td>
<td>0.80</td>
<td>0.66</td>
<td>0.52</td>
<td>0.36</td>
<td>0.47</td>
</tr>
<tr>
<td>PF-exclusions</td>
<td>0.78</td>
<td>0.64</td>
<td>0.48</td>
<td>0.39</td>
<td>0.46</td>
</tr>
</tbody>
</table>

POMS= Profile of Mood States; HADS= Hospital Anxiety and Depression Scale
Table 4: Responsiveness of the HOOS-PS, KOOS-PS, WOMAC Physical Function (PF) and the WOMAC items not included in the PS scales (PF-exclusions).

<table>
<thead>
<tr>
<th></th>
<th>Pre-surgery (mean, sd)</th>
<th>6 months post-surgery (mean, sd)</th>
<th>Standardized Response Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Hip Replacement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOOS-PS</td>
<td>55.9 (16.7)</td>
<td>25.4 (16.1)</td>
<td>1.5</td>
</tr>
<tr>
<td>PF</td>
<td>35.5 (12.0)</td>
<td>12.4 (10.9)</td>
<td>1.7</td>
</tr>
<tr>
<td>PF-exclusions</td>
<td>30.3 (9.6)</td>
<td>10.7 (9.1)</td>
<td>1.7</td>
</tr>
<tr>
<td><strong>Total Knee Replacement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KOOS-PS</td>
<td>55.3 (13.2)</td>
<td>36.9 (14.1)</td>
<td>1.4</td>
</tr>
<tr>
<td>PF</td>
<td>50.9 (18.0)</td>
<td>21.6 (16.6)</td>
<td>1.5</td>
</tr>
<tr>
<td>PF-exclusions</td>
<td>23.4 (8.1)</td>
<td>10.7 (9.1)</td>
<td>1.7</td>
</tr>
</tbody>
</table>