Cross-cultural adaptation, reliability and validity of the Turkish version of the Hospital for Special Surgery (HSS) Knee Score

Selnur NARİN¹, Bayram ÜNVER¹, Serkan BAKIRHAN², Özgür BOZAN³, Vasfi KARATOSUN⁴

¹Department of Orthopedic Physiotherapy, School of Physical Therapy and Rehabilitation, Dokuz Eylül University, Izmir, Turkey; ²Department of Physical Therapy and Rehabilitation, School of Health, Izmir University, Izmir, Turkey; ³School of Health Sciences, Gediz University, Izmir, Turkey; ⁴Department of Orthopedics and Traumatology, Dokuz Eylül University, Faculty of Medicine, Izmir, Turkey

Objective: The purpose of this study was to adapt the English version of the Hospital for Special Surgery (HSS) knee score for use in a Turkish population and to evaluate its validity, reliability and cultural adaptation.

Methods: Standard forward-back translation of the HSS knee score was performed and the Turkish version was applied in 73 patients. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), Mini-Mental State Examination and sit-to-stand test were also performed and analyzed. Internal consistency reliability was tested using Cronbach’s alpha. The intraclass correlation coefficient (ICC) was used to calculate the test-retest reliability at one-week intervals. Validity was assessed by calculating the Pearson correlation between the HSS, WOMAC and sit-to-stand test scores.

Results: The ICC ranged from 0.98 to 0.99 with high internal consistency (Cronbach’s alpha: 0.87). The WOMAC score correlated with total HSS score (r: -0.80, p<0.001) and sit-to-stand score (r: 0.12, p: 0.312).

Conclusion: The Turkish version of the HSS knee score is reliable and valid in evaluating the total knee arthroplasty in Turkish patients.

Key words: Reliability; The Hospital for Special Surgery knee score; translation; validity.

Osteoarthritis (OA) is one of the most prevalent chronic diseases worldwide and is associated with substantial impact on patients’ individual quality of life as well as on healthcare costs.¹ ¹² Patients with OA of the hip or knee joint experience pain, stiffness and loss of joint function. Total knee arthroplasty (TKA) is the most common and successful surgical treatment in cases in which conservative treatment does not relieve pain and improve function.³ ⁴ It is considered to be an effective intervention that improves patients’ quality of life, reduces pain and increases functional capability.⁵ ¹⁰

For clinicians, valid measurement tools provide important information to support effective clinical interpretation. For any test or measure of health status to be considered useful, it must be reliable, valid, responsive and acceptable.¹¹
Conventional approaches to the measurement of knee function have typically involved objectively defined parameters, such as radiological findings, strength, range of motion and ligamentous laxity. Different instruments for evaluating functional capacity and disability have been developed and used for specific knee conditions. The SF-36 and the Hospital for Special Surgery (HSS) knee rating scale are instruments widely used in TKA patients.

The HSS knee rating scale is a domain-specific questionnaire developed for use as a standardized instrument capable of measuring outcomes for patients with all knee disorders (such as OA and TKA) and designed specifically to evaluate patients with TKA of the knee. The instrument can be used to evaluate a patient before knee surgery and to monitor postsurgical function. The English version of the original index has been shown to be a reliable, valid, responsive and acceptable outcome measure. The HSS knee rating scale has been addressed in a number of research studies and is widely used by physiotherapists and orthopedic surgeons in clinical settings.

Most functional status questionnaires are constructed in English. Cross-cultural adaptation of validated outcome instruments has been advocated to facilitate their use in international, multicenter clinical trials. However, direct translation of questionnaires into other languages does not guarantee their validity. For measures to be used across cultures, the items must be not only translated linguistically but also adapted culturally to maintain its cross-culture content validity. To maintain the validity of the original instrument, while taking into consideration important cultural differences, a specific methodology has been developed for the adaptation process. Proper adaptation also serves to reduce the need for developing new instruments that have the same purpose.

The HSS knee rating scale is a commonly used in evaluating knee patients in Turkey as well as in Turks living outside of Turkey. For this reason, the purpose of this study was to cross-culturally adapt the HSS knee rating scale for Turkish patients and to determine the reliability and validity of the Turkish version of HSS knee scoring in patients with TKA.

**Patients and methods**

The translation and cross-cultural adaptation of the English version of the HSS knee rating scale into Turkish followed the recommended standard procedure. Two native Turkish speakers independently produced the forward translation of the HSS knee rating scale into the Turkish language. One (VK) was a medical doctor and the other (EBO) a university student. The translated text was then later translated back into English by a third party with fluent English.

To obtain a pre-final Turkish version of the questionnaire, an expert committee including translators, statisticians and health professionals compared the Turkish version with the original English version to detect errors of interpretation and nuances that might have been missed. The final stage of the adaptation process was to test the pre-final version. Pretesting of the pre-final Turkish version revealed no further difficulties with the questionnaire in 25 selected patients of different ages and social, ethnic, and educational backgrounds. The questionnaire was then approved by the committee and the developer of the HSS knee rating scale without any changes and was tested for validity and reliability on the study population.

The study was performed at the School of Physical Therapy and Rehabilitation and Department of Orthopedics and Traumatology in Faculty of Medicine of Dokuz Eylül University. The study included 73 patients who were independently mobile for a minimum of 6 months following TKA surgery. All operations were performed by the same surgeon (VK) using the paracapsular approach. The preoperative diagnosis was OA in all patients. Only patients with primary TKA were included in this study. The Mini-Mental State Examination (MMSE) was used in determining cooperation and only patients scoring over 20 points were included. Patients with rheumatoid arthritis, septic arthritis, gout, paresis, metastatic bone disease or previous fracture of the lower limbs were excluded from the study. Our study was approved by the University Ethics Committee and all participants were informed of the trial and signed written informed consent.

The HSS knee rating scale is a disease-specific test used to evaluate knee disabilities and methods of treatment, especially TKA. The HSS knee score gives a maximum of 100 points and considers subjective functional (52%) as well as objective examination criteria (48%). The domains include pain (30 points), function (22 points), range of motion (18 points), muscular strength (10 points), deformity (10 points), and instability (10 points). Subtractions are then made for walking aids, extension lag and varus-valgus deformity. The resultant score is classified in the ranges of >85 as ‘excellent’, 70 to 84 as ‘good’, 60 to 69 as ‘fair’ and <60 as ‘poor’. Deductions are made as follows: up to 3 points deducted for walking aids, up to 5 points for extension lag and 1 point for every 5° of varus-valgus deformity.
The MMSE was introduced by Folstein et al. in 1975. The Turkish version of the MMSE (translated by Güngen et al.) was used to evaluate the cognitive state of patients. Patients with moderate to severe cognitive impairment (score <20 on the MMSE) were not included in the study.

The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) is a frequently-used and recommended disease-specific questionnaire that has been found to be reliable and valid to determine self-report outcome after hip and knee replacement. The Turkish version of the WOMAC is considered a valid and reliable outcome measurement. For this reason, it was used to analyze the construct validity of the HSS. The Turkish version of the WOMAC consists of 24 questions in 3 subscales (5 pain, 2 stiffness, 17 physical function). It is available in Likert, visual analog scale and numeric rating scale formats. We used the numeric rating scale format in our study.

The ability to rise from a chair is an important activity of daily living as decreased ability can limit independence or contribute to institutionalization. The sit-to-stand test has been used as an indicator of postural control, fall risk, lower extremity strength and proprioception and as a measure of disability. The chair rise test is a more focused assessment of extensor mechanism function and is one of the activities used in functional indexes and in test batteries of physical functioning. The number of repetitions completed during a specified time interval (30 seconds) has been shown to be reliable and valid. The sit-to-stand test (30 seconds) is well suited to assess the physical function of those with lower limb arthritis, including those awaiting joint replacement surgery of the hip or knee.

Test-retest reliability measures the stability of a test over a given period of time. Intra-observer reliability was assessed through repeated assessment and additional tests were performed to assess validity. The most frequently assessed forms of reliability are test-retest reliability and internal consistency. The standard error of measurement (SEM), which represents the error between test and retest, was calculated for total scores. To determine test-retest reliability, HSS knee scoring was performed twice in one week (1 to 8 days). The intra-class correlation coefficient (ICC; range: 0.00 to 1.00) was used to determine test-retest reliability. A reliability coefficient between 0.60 and 0.80 is deemed very good and over 0.80 excellent. Internal consistency is an estimation of the strength of interrelated items in the test instrument and was assessed by calculating the Cronbach’s alpha, which ranges from 0 to 1, with higher values indicating higher internal consistency reliability. A linear regression was used for total score on test and retest to further assess reliability.

Validity is an index of how well a test measures what is supposed to measure. To evaluate the validity of the Turkish version of the HSS knee rating scale, the WOMAC, and sit-to-stand test were used and the Pearson correlation coefficient calculated. A value between 0 and 0.25 was considered no or poor correlation, 0.26 and 0.50 moderate, 0.51 and 0.75 good and greater than 0.75 very good.

<table>
<thead>
<tr>
<th>Patient characteristics</th>
<th>Number</th>
<th>Mean±SD</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td>71.56±7.74</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>8</td>
<td></td>
<td>11</td>
</tr>
<tr>
<td>Female</td>
<td>65</td>
<td></td>
<td>89</td>
</tr>
<tr>
<td>Employment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Housewife</td>
<td>55</td>
<td></td>
<td>75.3</td>
</tr>
<tr>
<td>Working</td>
<td>1</td>
<td></td>
<td>1.4</td>
</tr>
<tr>
<td>Retired</td>
<td>17</td>
<td></td>
<td>23.3</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary school</td>
<td>42</td>
<td></td>
<td>57.5</td>
</tr>
<tr>
<td>High school</td>
<td>27</td>
<td></td>
<td>37.0</td>
</tr>
<tr>
<td>University</td>
<td>4</td>
<td></td>
<td>5.5</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td></td>
<td>80.32±16.60</td>
<td></td>
</tr>
<tr>
<td>Height (cm)</td>
<td></td>
<td>158.91±7.55</td>
<td></td>
</tr>
<tr>
<td>Body Mass Index</td>
<td></td>
<td>32.84±6.60</td>
<td></td>
</tr>
<tr>
<td>Duration of operation (years)</td>
<td>5.18±2.60</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SPPS for Windows v.15.0 (SPSS Inc., Chicago, IL, USA) was used for all statistical analyses. Data were tested for normality using the Kolmogorov-Smirnov test. The measured variables are presented as mean±SD while the categorical variables are expressed as percentages. Validity was assessed by calculating the Pearson correlation coefficients between the HSS and other scales.

**Results**

All 73 patients filled in the HSS questionnaires and were investigated clinically. Participants’ sociodemographic and clinical data are presented in Table 1. Patients did not report difficulty in understanding and completing the Turkish version of the HSS questionnaire.

Intraclass correlation coefficients ranged from 0.98 to 0.99 (Table 2). The computed SEM values were low (range: 1.07 to 1.84), supporting the reliability values obtained, which reflect the individual change above the measurement error and indicate minimal clinical change. These observations also are supported by the linear regression and the scatterplot (regression coefficient: 0.99) of total scores in two consecutive measurements (Fig. 1). This indicates a strong relationship between the data collected on both occasions. There were no differences between the means of test-retest values with respect to the subscores and total scores. As for internal consistency, the Turkish version of the HSS scale had a Cronbach’s alpha of 0.87 when all items were considered.

The Turkish version of the HSS scale appeared to be valid (ICC range: 0.98 to 0.99) as assessed using the WOMAC and sit-to-stand test scores (Table 3). There was a significant negative correlation between the WOMAC score and total HSS score (r=-0.80, p<0.001) and no significant correlation between WOMAC and sit-to-stand scores (r=0.12, p=0.312). The WOMAC score had the highest correlation with the total score of the HSS scale. The sit-to-stand test had a poor correlation with the total score (r=0.12, p=0.312) (Table 3).

**Discussion**

The Turkish version of the HSS questionnaire was translated according to international standardized guidelines for patients who underwent TKA surgery. The agreement, reliability and validity of the Turkish version of the HSS questionnaire are acceptable in Turkish-speaking patients with TKA.

The HSS knee rating scale was developed by Insall et al. in 1976 to evaluate the pre- and postoperative assessment of four different types of knee replacement.

---

**Table 2.** Reliability (test-retest) values of HSS questionnaire.

<table>
<thead>
<tr>
<th>HSS score</th>
<th>Test'</th>
<th>Retest'</th>
<th>ICC†</th>
<th>SEM†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain</td>
<td>24.58 (6.35)</td>
<td>24.65 (6.37)</td>
<td>0.999 (0.998-0.999)</td>
<td>0.74 (0.68-0.80)</td>
</tr>
<tr>
<td>Function</td>
<td>15.21 (4.04)</td>
<td>15.24 (4.03)</td>
<td>0.999 (0.998-0.999)</td>
<td>0.47 (0.33-0.61)</td>
</tr>
<tr>
<td>ROM</td>
<td>11.28 (1.45)</td>
<td>11.28 (1.47)</td>
<td>0.990 (0.985-0.994)</td>
<td>0.71 (0.70-1.72)</td>
</tr>
<tr>
<td>Muscle strength</td>
<td>8.50 (1.51)</td>
<td>8.50 (1.51)</td>
<td>0.997 (0.996-0.998)</td>
<td>0.17 (0.09-0.25)</td>
</tr>
<tr>
<td>Deformation</td>
<td>9.93 (0.58)</td>
<td>9.93 (0.58)</td>
<td>0.999 (0.999-0.999)</td>
<td>0.00 (0.00-0.00)</td>
</tr>
<tr>
<td>Instability</td>
<td>10.00 (0.00)</td>
<td>10.00 (0.00)</td>
<td>0.999 (0.999-0.999)</td>
<td>0.00 (0.00-0.00)</td>
</tr>
<tr>
<td>Total</td>
<td>79.27 (10.25)</td>
<td>79.35 (10.25)</td>
<td>0.999 (0.998-0.999)</td>
<td>1.20 (1.12-1.28)</td>
</tr>
</tbody>
</table>

*Data are presented as means, with standard deviations in parentheses. †Results were calculated with 95% confidence intervals in parentheses. ‡The correlation and t values cannot be computed as the standard error of the difference is 0. HSS: The Hospital for Special Surgery; ICC: Intraclass correlation coefficient; ROM: Range of motion; SEM: Standard error of measurement.

**Table 3.** Pearson correlation (r values and p values) between HSS and WOMAC, sit-to-stand scores.

<table>
<thead>
<tr>
<th>Test</th>
<th>HSS (total)</th>
<th>Sit-to-stand</th>
<th>WOMAC</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSS (total)</td>
<td>r</td>
<td>1.00</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>0.312</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sit-to-stand</td>
<td>r</td>
<td>0.12</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>0.312</td>
<td></td>
</tr>
<tr>
<td>WOMAC (total)</td>
<td>r</td>
<td>-0.80*</td>
<td>-0.02</td>
</tr>
<tr>
<td></td>
<td>p</td>
<td>&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.001 level (2-tailed). Significant values for r and p are written in bold. HSS: The Hospital for Special Surgery; WOMAC: The Western Ontario and McMaster Universities Osteoarthritis Index.
prosthesis and has been shown to be reliable and valid in these patients. While different tools have been used in the evaluation of arthroplasty and knee disorders in other countries and cultural adaptations, the HSS had not previously been used in these disorders. This translation and cultural adaptation was conducted in Turkish and intended for use both in Turkey and in countries with a Turkish population. The Turkish population in European Union countries currently stands at 3 million.

In general use, the HSS knee scoring has a wide following and is deemed easy to use and quick to record. A recent study of the inter-observer reproducibility of a number of knee scoring systems found that while the HSS knee score had good overall inter-observer correlation coefficients, the reproducibility of some parts of the score was poor. The WOMAC, SF-36 and Oxford Knee Scores have undergone the most thorough assessment of reliability and validity and are therefore the most appropriate for the assessment of outcome after TKA.

The Turkish version of the HSS questionnaire was fully filled out and easily completed by all patients. The short time required to complete the questionnaire suggested that the Turkish version of the HSS questionnaire is well comprehended by native Turkish patients. In addition, the short time needed to evaluate the questionnaire by researchers confirms that the Turkish HSS questionnaire is a practical and easily assessed tool.

In many studies, a sample size of 30 subjects is not adequate for studying the reliability or validity of the instrument. When statistical estimates are derived from small samples, confidence intervals are wide and reflect a high degree of uncertainty in the precision of the reliability coefficient. Terwee et al. suggested that a sample size of at least 50 subjects should be used. Therefore, the number of patients in the current study (n=73) was considered sufficient to conduct validity and reliability analysis.

In an independent study on a different knee system, HSS scores at the 6th postoperative month were 75.6 points. Other studies reported points ranging from 85.0 to 93.0 at the one-year follow-up. Similar to the literature, at a mean time of 5.19 years after surgery, mean HSS scores were 79.35. As it is well known that maximum functional gain is achieved in the first six months following TKA surgery, patients with a minimum of six months of follow-up were included in this study.

Reliability is classically considered to comprise two domains: test-retest repeatability and internal consistency. Although ICCs have been commonly used as a measure of reliability (test-retest) in previous validations of various knee instruments, a more appropriate approach is to calculate the 95% limits of agreement for the differences between the two repeated measurements. This analysis showed acceptable agreement for the HSS scores. Additionally, the HSS scores showed acceptable internal consistency (with Cronbach alpha values clearly above 0.70). We found no differences in the means of test-retest subgroups when pain, function, ROM, muscle strength, deformation, instability and total scores were considered (Table 2).

Table 4. Descriptive information for functional tests and MMSE.

<table>
<thead>
<tr>
<th>Test</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSS (total)</td>
<td>79.35 (10.25)</td>
</tr>
<tr>
<td>WOMAC (total)</td>
<td>15.27 (14.81)</td>
</tr>
<tr>
<td>Sit-to-stand</td>
<td>12.23 (2.54)</td>
</tr>
<tr>
<td>MMSE</td>
<td>23.15 (2.63)</td>
</tr>
</tbody>
</table>

*Data are presented as means, with standard deviations in parentheses. HSS: The Hospital for Special Surgery; MMSE: The mini-mental state examination; WOMAC: The Western Ontario and McMaster Universities Osteoarthritis Index.
for all items of the questionnaire. Criterion validity, or instrumental validity, is used to demonstrate the accuracy of the instrument by comparing it with the truth or another instrument that has been proven to be valid.[54]

The WOMAC score and sit-to-stand test were used to determine criterion validity.

WOMAC scores were highly correlated and the sit-to-stand test scores poorly correlated with the total HSS scores (Table 4). The sit-to-stand test did not have enough sensitivity and is therefore not recommended for use alone in assessing functional capacity. The Knee Outcome Survey-Activities of Daily Living Scale (KOS-ADLS) which has been adapted to Turkish[55] could be useful for this purpose in future studies.

In conclusion, the HSS knee rating scale was successfully translated and adapted into the Turkish language. The Turkish version of the HSS scale met the criteria of reliability and validity in measuring symptoms and functional limitations in patients with knee pain.

Acknowledgment: The authors are grateful to Prof. Dr. Hülya ELLIDOKUZ for statistical suggestions and to Elif Bike ÖSÜN for her help in translation.

Conflicts of Interest: No conflicts declared.

References
36. Segura-Ortí E, Martinez-Olmos FJ. Test-retest reliability and minimal detectable change scores for sit-to-stand-to-sit tests, the six-minute walk test, the one-leg heel-rise test, and handgrip strength in people undergoing hemodialysis. Phys Ther 2011;91:1244-52. CrossRef


