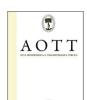


Contents lists available at ScienceDirect

## Acta Orthopaedica et Traumatologica Turcica

journal homepage: https://www.elsevier.com/locate/aott



# Evidence for reliability, validity and responsiveness of Turkish version of Hip Outcome Score<sup>★</sup>



Gökhan Polat <sup>a, \*</sup>, Derya Çelik <sup>b</sup>, Hilal Çil <sup>b</sup>, Mehmet Erdil <sup>c</sup>, Mehmet Aşık <sup>a</sup>

- <sup>a</sup> Istanbul University, Istanbul Medical Faculty, Department of Orthopaedics and Traumatology, Istanbul, Turkey
- <sup>b</sup> Istanbul University, Faculty of Health Science, Division of Physiotherapy and Rehabilitation, Istanbul, Turkey
- <sup>c</sup> Istanbul Medipol University, Department of Orthopaedics and Traumatology, Istanbul, Turkey

#### ARTICLE INFO

Article history: Received 5 October 2016 Received in revised form 19 March 2017 Accepted 15 May 2017 Available online 9 June 2017

Keywords:
Hip Outcome Score
Turkish
Clinometric
Outcome measurement
Translation
Hip arthroscopy

#### ABSTRACT

*Background:* Hip Outcome Score (HOS), originally developed in English, assesses the severity of hip pathology. To date, no Turkish version of the questionnaire exists.

*Purpose*: The aim of our study was to translate the HOS into Turkish and verify its psychometric properties.

Methods: The translation and cultural adaptation were performed according to international recommendations in five stages: The HOS was translated into Turkish, consistent with published methodological guidelines. The process included 2 forward translations, followed by the synthesis of these translations, and 2 backward translations, followed by an analysis of the translations and creation of the final version. The measurement properties of the Turkish HOS (internal consistency, construct validity, floor and ceiling effects and responsiveness) were tested in 130 patients.

Results: A committee consisting of the four translators agreed with the final version of the HOS (HOS-Tr). The internal consistency and the test-retest reliability of the HOS-Tr-ADL and HOS-Tr-S subscales were excellent. Correlations between the HOS-Tr and convergent validity of the with HHS and NAHS were fair to good. The responsiveness of the HOS-Tr-ADL and HOS-Tr-S subscales were 3.4 to 1.4 for patients treated with surgically and 0.9 to 1.1 for patients treated with non-surgically.

*Conclusion:* The HOS-Tr is understandable, reliably, valid, and responsive for Turkish-speaking patients with hip pathology.

Level of Evidence: Level 3 Diagnostic Study.

© 2017 Turkish Association of Orthopaedics and Traumatology. Publishing services by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/

4.0/).

## Introduction

Patient-reported outcomes (PROs) provide insights from the patient's perspective of the impact of disease and are effective tools for the evaluation of the treatment results for surgeons. Many PROs

have been developed for the evaluation of hip surgeries including Hip Outcome Score (HOS), Non-Arthritic Hip Score (NAHS), Oxford Hip Score (OHS), Hip Disability and Osteoarthritis Outcome (HOOS), Western Ontario and McMaster Universities Osteoarthritis Index, (WOMAC), International Hip Outcome Tool-33. <sup>1–5</sup> Of these, HOS was designed to measure not only the functional impairment of the patients in daily living (HOS-ADL) but also the functional impairment of the patients in sportive activities (HOS-S) including many specific movements that may push the limits of hip joint functions. <sup>6–8</sup>

Before using PROs in a society other than that in which the outcome measure was developed, it should be translated and culturally adapted. The PROs that have been translated into Turkish and psychometrically tested only include HHS, WOMAC, OHS and HOOS -Physical Function Short-Form. 9–13

 $<sup>^{\</sup>star}$  Institutional Review Board has approved this study. (Istanbul University IRB (2016/255)).

<sup>\*</sup> Corresponding author. Istanbul University, Istanbul Faculty of Medicine, Department of Orthopaedics and Traumatology, Capa Fatih Istanbul 34093, Turkey. Fax: +90 212 635 12 36.

E-mail addresses: gokhanpolat7@gmail.com (G. Polat), ptderya@hotmail.com (D. Çelik), cilhilal@gmail.com (H. Çil), drmehmeterdil@gmail.com (M. Erdil), mehmetasik@hotmail.com (M. Aşık).

Peer review under responsibility of Turkish Association of Orthopaedics and Traumatology.

The aim of this study was to translate and adapt the HOS questionnaire into Turkish and to test the psychometric properties of the HOS in terms of reliability, validity, and responsiveness.

#### Materials and methods

#### Translation and cross-cultural adaptation

Translation and cross-cultural adaptation of the HOS was performed in 5 stages, consistent with the stages recommended by Beaton et al.<sup>14</sup> In the first stage, 2 Turkish individuals with a good command of English were responsible for the literal and conceptual translation of the HOS Form. The informed translator was a physical therapist, and the uninformed translator was a translator and interpreter both spoke Turkish as their mother tongue. In the second stage, both translations were compared and reviewed by a bilingual individual who highlighted any conceptual errors or inconsistencies in the translations to establish the first Turkish translation. In the third stage, after the first Turkish translation was agreed upon, 2 native English speakers with a good command of Turkish separately translated the finalized Turkish translation back into English. Both translators were unaware of the purpose of the study and had no access to the original English version. In the fourth stage, the back translated version of the HOS was compared to the initial English version of the HOS by a committee consisting of the four translators. After discussing the discrepancies, the committee finalized and approved the Turkish version of the HOS Form (HOS-Tr). In the final stage, preliminary testing was performed to determine comprehension of the Turkish version (Appendix).

#### Patients reported outcomes

HOS-ADL includes 19 questions that 17 of which are scored and was designed to measure the functional status during daily living activities. The second part of the questionnaire called HOS-S that includes 9 questions related with sports activities like running, jumping etc. The highest potential of HOS-ADL is 68 and HOS-S is 36. This value is then multiplied by 100 to get a percentage. HHS is a well-known region specific outcome measure used by clinicians to measure pain, function and range of motion of the hip joint. NAHS is also a disease specific outcome measure for hip joint that measures the pain and functional limitations during the last 48 h. 4

## **Participants**

This study was approved by the Institutional Review Board (2016/255) and an informed consent form was signed by all participants. The study was performed between January 2015 and December 2015. The eligibility criteria were (1) 18–60 years of age, (2) hip pathology including acetabular dysplasia, labral tears, FAI, tendon or muscle injuries, (3) patients who had treated surgically via hip arthroscopy (4) ability to read and write in Turkish. Patients who had Tonnis grade 3 and 4 degenerative arthritis, who had previous or additional lower extremity surgeries that may affect the functional evaluation, patients who did not perform any sports and who did not want to attend the study, were excluded. Diagnoses were established by 2 orthopedic surgeons. Age, gender, occupations, involved side and diagnosis of the participants were recorded.

One hundred thirty consecutive patients with a variety of hip disorders were invited to complete the HOS-Tr and the Turkish version of the HHS and NAHS. Subgroups of 30 patients were asked to complete the HOS-Tr again 7–14 days after their first completion to determine the test-retest reliability. To minimize the risk of short-term clinical change, no treatment was provided during this period. Responsiveness was assessed in 100 patients who were surgically treated and 30 patients who were treated non-surgically.

## Preliminary testing

Preliminary testing was conducted on 30 of the 130 patients (11 males, mean age 32.8  $\pm$  10.6 (range 21–54)) who fulfilled the eligibility criteria of the study to determine comprehension of the Turkish version. Following completion of the questionnaire by each patient, two researchers performed an interview in which the patients were asked if they had any difficulties understanding the questions. The questions that were difficult to understand were noted, and the patients were asked for their recommendations for revisions.

#### Statistical analysis

All statistical analyses were performed with the Statistical Package for the Social Sciences 20.0 (SPSS Inc, Chicago, IL, USA). The level of significance was set at  $p \leq 0.05$ . Descriptive statistics were calculated for all variables. These included frequency counts, the percentage for nominal variables, measures of central tendency (means and medians) and dispersion (standard deviations and ranges) for continuous variables. Before the statistical analysis, the Shapiro Wilk test was used to test for normal distribution of data. Dependent variables were compared using an analysis of variance for repeated measures. The measurement properties analyzed in this study for the instruments included internal consistency, the test-retest reliability, agreement, construct validity, ceiling and floor effects and responsiveness.

## Internal consistency

Internal consistency was used to determine the interrelatedness among the items of the HOS-Tr. An inter-item correlation matrix was used to indicate whether one of the items did not correlate positively with the other items. A Cronbach alpha value ranging from 0.70 to 0.95 was considered to be adequate. <sup>16</sup> Data from the patients included in the first administration of the HOS-Tr were used to assess internal consistency.

## Test-retest reliability

Test-retest reliability represents a scale's ability to yield consistent results when administered on separate occasions during a period when an individual's status has remained stable. <sup>17</sup> Intraclass correlation coefficients (ICCs) were calculated using a 2-way, mixed-model under consistency.

## Agreement

Agreement was assessed with the standard error of measurement (SEM) and minimal detectable change (MDC). The ICC was used to calculate the SEM, which is an index of measurement precision. The SEM is calculated as the SD of the scores

time the square root of (1-ICC). The minimal detectable change (MDC) refers to the minimal amount of change that is within the measurement error. The SEM was used to determine the minimum detectable change at the 95% limits of confidence (MDC95%) and was calculated as the SEM times 1.96 time the square root of  $2.^{16}$ 

#### Validity

Evidence for construct validity of the HOS-Tr was provided by determining its relationship with HHS and NAHS. Content validity was assessed by the distribution of the scores and occurrence of ceiling and floor effects. Floor and ceiling effects of the HOS-Tr at the first and second assessment were assessed by calculating the proportion of patients scoring the minimum or maximum values on the scale relative to the total number of patients. We considered scores between 0% and 10% to be minimum scores and scores between 90% and 100% to be maximum scores. Floor and ceiling effects were considered to be relevant if greater than 30% of the patients had a score at the limits of the scale. <sup>18</sup>

## Responsiveness

Responsiveness was assessed in 100 patients who were treated by surgically and 30 patients who was treated by conservatively. Effect sizes (ES) were determined by calculating the differences in the means of baseline and follow-up data divided by the standard deviation at baseline demonstrated. <sup>19</sup>

#### Results

## Translation and cultural adaptation

During the translation process the translators had difficulty in translating 3 words; "landing," "cutting/lateral movements" "stepping-up and down curbs." A consensus was reached on the translation so that the meaning of the questions did not change. The distance unit had to be changed to metric units. "Running one mile" appears in the original HOS was changed "running to 1–2 km". However, the patients felt more comfortable explaining distance as minutes spent walking. Therefore, we included both distance and duration in the questionnaire. The preliminary testing did not show any difficulty in patients' understanding of these words. In the assessment of daily living activities, some patients needed to inform the researchers regarding that they were not using a bath tub in their home. So the patients were asked to simulate this activity with trying to step in a bath tub that needs a deep hip flexion and rotation of the hip joint and answer according to this activity. In assessment of sports activities, some of the patients needed to informed the researchers regarding they were not playing golf. These patients were asked to simulate this activity with a long stick that needs hyperextension and rotation of the hip joint.

## Measurement properties and testing

The demographics and clinical characteristics of the participants were presented in Table 1. 130 patients completed all of the questionnaires at the first assessment by themselves in a room under custody of the researchers. Comprehensibility and acceptance of the questionnaire determined by the ratio of

**Table 1** Patient demographics (n = 130).

Characteristic	Value
Age, mean (SD)	34.8 (10.6)
Male gender, n (%)	64
Occupation n(%)	
Housewife	21 (16.1)
Retired	13 (10.0)
Labor	45 (34.6)
Whitecollor	29 (22.3)
Student	14 (10.7)
Athletes	8 (6.1)
Involved side n (%)	
Right leg	59 (45.4)
Diagnosis n (%)	
Labral Tear	25 (19.2)
Labral Tear + Acetabular Dysplasia	10 (7.6)
Acetabular Dysplasia + Chondropathy	2 (1.5)
Labral Tear + FAI	74 (56.9)
Extraarticular	11 (8.4)
Osteoid Osteoma	2 (1.5)
Avascular Necrosis of Femoral Head	2 (1.5)
Synovial Mass	3 (2.3)

unanswered questions were good since there were no unanswered questions. Thirty of the 130 participants who were given an appointment for nonsurgical treatment included for the test-retest assessment.

## Reliability

The internal consistency of the first assessment of the HOS-Tr-ADL and HOS-Tr-S for were strong, with a Cronbach's  $\alpha$  value of 0.95 (95% CI, 0.94–0.97) and 0.91 (95% CI, 0.90–0.91). The interitem correlation matrix did not show any low or negative interitem correlation. The interval between the first and second assessments was 8.2 days. The test-retest reliability was 0.98 (95% CI, 0.97–0.99) and 0.97 (95% CI, 0.96–0.99) for ADL and Sports subscales. The results of internal consistency, the test-retest reliability and comparisons with other translated versions of the HOS are provided in Table 2.

## Agreement

The SEM and MDC were determined to be 1.6 and 4.3 for HOS-Tr-ADL, 0.96 and 2.6 for HOS-Tr-S.

**Table 2**Reliability of the HOS, including the Turkish version.

	Language versions	Test-retest reliability (ICC) HOS-Tr-ADL (n = 30) HOS- Tr-S (n = 30)		Cronbach's Alpha HOS-Tr-ADL (n = 130) HOS- Tr-S (n = 130)	
Martin	English	0.98	0.92	_	_
Lee	Kore	0.95	0.929	>0.90	>0.90
de Oliveira	Portuguese	_	-	_	_
Naal	German	0.94	0.89	>0.90	>0.90
Seijas	Spanish	0.95	0.94	>0.90	>0.90
Present study	Turkish	0.98	0.97	0.95	0.91

Abbreviation: ICC, Intraclass Correlation Coefficient; HOS, Hip Outcome Score.

#### Validity

The HOS-Tr-ADL and HOS-Tr-S subscales demonstrated very good correlation with the HHS ( $r=0.56\ p=0.001,\ 0.25\ p=0.003$  respectively) and fair correlation with NAHS ( $r=0.21\ p=0.01,\ 0.33\ p=0.001$  respectively).

## Floor and ceiling effects

Floor and ceiling effects and the number of items answered were identical during the test and retest examinations for both HOS-Tr-ADL and HOS-Tr-S subscales. Ceiling effect was observed in 2% of patients of the HOS-Tr-ADL subscale whereas floor effect was not observed.

## Responsiveness

In the surgical treatment group, baseline assessment on the HOS-Tr was compared with the post-op HOS-Tr at 1-year follow-up with 100 patients (54 males; mean  $\pm$  SD age, 36.2  $\pm$  8.4 range, 30–59 years). The mean and standard deviation of the baseline, and 1 year follow-up values of the HOS-Tr-ADL and HOS-Tr-S subscales were 47.1  $\pm$  6.01, 67.4  $\pm$  6.9 and 22.2  $\pm$  4.1, 28.0  $\pm$  4.3 respectively. The subscales indicated a large effect size at 1 year follow-up ES of 95% CI:3.4 and 1.4 respectively. In the nonsurgical treatment group, baseline assessment on the HOS-Tr was compared with 3 months' follow-up of HOS-Tr for 30 patients (10 males; mean  $\pm$  SD age, 35.4  $\pm$  7.2 range, 30–49 years). The mean and SD of the HOS-Tr-ADL and HOS-Tr-S in nonsurgical treatment group were 57.2  $\pm$  7.4, 64.1  $\pm$  7.5 and 22.1  $\pm$  3.4, 25.8  $\pm$  4.6 respectively. The ES was found 0.9 and 1.1 respectively on HOS-Tr-ADL and HOS-Tr-S (Table 3).

## Discussion

The aim of this study was to translate and culturally adapt the HOS into Turkish and provide reliability, validity and responsiveness for the translated version based on a sample of Turkish-speaking patients with hip injuries.

We acknowledge certain limitations of our study. Patients were not very compliant to complete the retest assessment therefore only 23% percent of the patients completed the second assessment. Therefore, the sample size was low for the reliability analysis, which reduced the precision of our estimates. We only assessed the convergent validity of HOS-Tr but divergent validity was not performed. Nevertheless, minimal clinically important

**Table 3**Responsiveness Turkish version of the HOS.

Measurements	Mean ± SD		ES
Surgical Treatment (n = 100)	Baseline	1 year follow-up	
HOS-ADL	$47.1 \pm 6.0$	$67.4 \pm 6.9$	3.4
HOS-S	$22.2 \pm 4.1$	$28.0 \pm 4.3$	1.4
Conservative Treatment $(n = 30)$	Baseline	3 months follow-up	
HOS-ADL	$57.2 \pm 7.4$	$64.1 \pm 7.5$	0.9
HOS-S	$22.1 \pm 3.4$	$25.8 \pm 4.6$	1.1

differences in patients with various hip pathologies should be assessed.

Internal consistency of the Turkish version, using Cronbach alpha, was 0.95 for HOS-Tr-ADL and 0.91 for HOS-Tr-S which is considered excellent and higher values previously reported in the literature.  $^{1,20-22}$  Test-retest reliability of the HOS-Tr-ADL and HOS-Tr-S subscales were found excellent (ICC =0.98, ICC =0.97 respectively) for such as original version (ICC =0.98, ICC =0.92) and similar to other Korean (CC =0.98, ICC =0.97), German (ICC =0.94, ICC =0.94) and Spanish (ICC =0.95, ICC =0.94) versions.  $^{20-22}$ 

The present study provides support for the construct validity of the scale, comparing HOS-Tr and HHS and NAHS of the Turkish version. The correlation coefficient with HOS-Tr-ADL and HOS-Tr-S and Turkish version of the NAHS were fair to good (r = 0.21, r = 0.33). The highest value was found between HOS-Tr-ADL and HHS (r = 0.56). Naal et al reported the weak correlation coefficients with the Mental Component Scale of Short Form 12 (r = -0.08) and excellent correlation with WOMAC function subscale (r=-0.90) and German version of the HOS. <sup>21</sup> Spanish HOS was correlated with the WOMAC subscales and found good to very good correlation (r = 0.49 to 0.77).<sup>22</sup> Martin et al showed a strong correlation between HOS and the SF-36 physical function and physical component subscale 0.76 and 0.74 respectively for the HOS-ADL subscale and 0.72 and 0.68 for the HOS-sports subscale as expected the correlation with the SF-36 mental components was weaker.<sup>6</sup> The Korean version of the HOS-ADL and HOS-S subscales showed poor to good correlation ( $rh_0 = 0.12$  to 0.68) with SF-36 subscales and good to very good correlation  $(rho = 0.38 \text{ to } 0.78) \text{ with HOOS subscales and total HOOS scores.}^{10}$ In the present study, we did not use SF-36 for convergent and divergent validity therefore, we could not compare our validity result with literature.

2% of the patients scored or maximum score but it was still below %30 indicating that floor effect. Martin et al reported only one patient who scored 100 point for both subscales.<sup>6</sup> In the German version, ceiling effect was higher than the floor effect in the HOS-ADL and the HOS-S subscales.<sup>21</sup> Spanish version of the HOS showed ceiling effect was observed in 6% and 12% for ADL and sports subscale, respectively. Floor effect was found in 3% and 37% ADL and sports subscale, respectively. No floor or ceiling effect was observed also in Korean version of the HOS.<sup>20</sup>

Responsiveness, based on the completion of the HOS-Tr prior to and 1 year follow-up for surgical treatment group showed larger ES compare to nonsurgical treatment group which were followed at 3 months. This is because the patients may provide a better improvement with surgery. The only study presented responsiveness was the Korean version of the HOS, however, the responsiveness was determined by using Wilcoxon signed-rank test. Therefore, we could not compare our results with literature.

In conclusion the HOS-Tr provides strong evidence that the HOS-Tr has sufficient reliability, validity, and responsiveness, with values similar to those reported for the original and other translated versions.

#### **Disclosure**

No funding was received by none of the authors related to this study.

## **Appendix**

İsim: Tarih:

## KALÇA DEĞERLENDİRME SKORU (HOS)

## Günlük Yaşam Aktivite Ölçeği.

Lütfen <u>her soruyu</u> cevaplarken geçtiğimiz hafta boyunca durumunuzu en iyi açıklayan <u>tek seçeneği</u> işaretleyiniz. Soruda tanımlanan aktiviteler kalçanızdan değil de vücudunuzun başka bir bölgesi tarafından kısıtlanıyorsa uygulanamaz kısmını işaretleyin.

	Hiç zor değil	Biraz zor	Orta derecede zor	Çok zor	İmkânsız	Uygulanamaz
15 dakika boyunca ayakta durmak						
Arabaya inip binmek						
Dik yokuş çıkmak						
Dik yokuş inmek						
1 kat merdiven çıkmak						
1 kat merdiven inmek						
Kaldırıma çıkıp inmek						
Çömelmek						
Küvete girip çıkmak						
Yürümeye başlamak						
Yaklaşık 10 dakika boyunca yürümek						
15 dakika veya daha fazla yürümek						

## Aşağıdaki faaliyetleri yaparken kalçanızdan dolayı ne kadar zorluk çekiyorsunuz?

İmkânsız	Uygulanamaz
	Imkānsiz

Kalça probleminiz ortaya çıkmadan önceki iş yapabilme seviyenizin 100, günlük aktivitelerinizin hiçbirini yerine getiremediğiniz seviyenin 0 olduğunu varsayarsanız, günlük aktiviteleri yerine getirme seviyeniz için 0 ila 100 arasında kaç puan verirdiniz.

Puan verilmemiş

	Hiç zor değil	Biraz zor	Orta derecede zor	Çok zor	İmkânsız	Uygulanamaz
Çorap ve ayakkabı giymek						
15 dakika boyunca oturmak						

## KALÇA DEĞERLENDİRME SKORU (HOS) Spor Ölçeği.

## Aşağıdaki aktiviteleri yaparken kalçanızdan dolayı ne kadar zorluk çekiyorsunuz?

	Hiç zor değil	Biraz zor	Orta derecede zor	Çok zor	İmkânsız	Uygulanamaz
1,5 kilometre (20 dakika) koşmak						
Zıplamak						
Golf sopası gibi cisimleri savurmak						
Sıçrama sonrasında yere inmek						
Aniden hareketlenmek ve durmak						
Yana koşular sırasında aniden durmak						
Hızlı yürüyüş gibi düşük etkili aktiviteler						
Alıştığınız şekilde aktivite yapabilme kabiliyeti						
İstediğiniz sürece, istediğiniz spor aktivitesini yapabilme kabiliyeti						

Kalça probleminiz ortaya çıkmadan önceki iş yapabilme seviyenizin 100, günlük aktivitelerinizin hiçbirini yerine getiremediğiniz seviyenin 0 olduğunu varsayarsanız, spor aktiviteleri yerine getirme seviyeniz için 0 ila 100 arasında kaç puan verirdiniz. %.....

Şu anki iş yapabilme seviyenizin nasıl olduğunu düsünüvorsunuz?

auşunuyorsunuz.						
Normal	Neredeyse normal	Normal değil	Hiç normal değil			

#### References

- Bellamy N. The WOMAC knee and hip osteoarthritis indices: development, validation, globalization and influence on the development of the AUSCAN hand osteoarthritis indices. Clin Exp Rheumatol. 2005;23(5 Suppl 39): 5148–5153.
- Christensen CP, Althausen PL, Mittleman MA, Lee JA, McCarthy JC. The nonarthritic hip score: reliable and validated. Clin Orthop Relat Res. 2003 Jan;406(1):75–83.
- 3. Dawson J, Fitzpatrick R, Frost S, Gundle R, McLardy-Smith P, Murray D. Evidence for the validity of a patient-based instrument for assessment of outcome after revision hip replacement. *J Bone Joint Surg Br.* 2001 Nov;83(8): 1125–1129.
- Klässbo M, Larsson E, Mannevik E. Hip disability and osteoarthritis outcome score. An extension of the western Ontario and McMaster Universities osteoarthritis index. Scand J Rheumatol. 2003;32(1):46–51.
- Mohtadi NG, Griffin DR, Pedersen ME, et al. Multicenter Arthroscopy of the Hip Outcomes Research Network. The Development and validation of a selfadministered quality-of-life outcome measure for young, active patients with symptomatic hip disease: the International Hip Outcome Tool (iHOT-33). Arthroscopy. 2012;28(5):595–605.
- Martin RL, Kelly BT, Philippon MJ. Evidence of validity for the hip outcome score. Arthroscopy. 2006;22(12):1304–1311.
- Martin RL, Philippon MJ. Evidence of validity for the hip outcome score in hip arthroscopy. Arthroscopy. 2007;23(8):822–826.
- 8. Martin RL, Philippon MJ. Evidence of reliability and responsiveness for the hip outcome score. *Arthroscopy*. 2008;24(6):676–682.
- Celik D, Can C, Aslan Y, Ceylan HH, Bilsel K, Ozdincler AR. Translation, crosscultural adaptation, and validation of the Turkish version of the Harris Hip Score. *Hip Int*. 2014;24(5):473–479.
- Tüzün EH, Eker L, Aytar A, Daşkapan A, Bayramoğlu M. Acceptability, reliability, validity and responsiveness of the Turkish version of WOMAC osteoarthritis index. Osteoarthr Cartil. 2005 Jan;13(1):28–33.

- Tuğay BU, Tuğay N, Güney H, Hazar Z, Yüksel İ, Atilla B. Cross-cultural adaptation and validation of the Turkish version of Oxford hip score. *Arch Orthop Trauma Surg.* 2015 [un;135(6):879–889.
- 12. Yilmaz O, Gul ED, Bodur H. Cross-cultural adaptation and validation of the Turkish version of the hip disability and osteoarthritis outcome score-physical function short-form (HOOS-PS). *Rheumatol Int.* 2014 Jan;34(1):43–49.
- Rolfson O, Eresian Chenok K, Bohm E, et al. Patient-reported outcome measures working group of the international society of arthroplasty registries. Patientreported outcome measures in arthroplasty registries. *Acta Orthop.* 2016 Jul;87(Suppl 1):3–8.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)*. 2000;25(24):3186–3191.
- **15.** Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. An end-result study using a new method of result evaluation. *J Bone Joint Surg Am.* 1969;51(4):737–755.
- Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(2):34–42.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174.
- 18. Nunnally JC, Bernstein IH. *Psychometric Theory*. 3rd ed. New York, NY: McGraw Hill; 1994.
- 19. De Vet HC, Terwee CB, Bouter LM. Current challenges in clinimetrics. *J Clin Epidemiol.* 2003;56(12):1137–1141.
- Lee YK, Ha YC, Martin RL, Hwang DS, Koo KH. Transcultural adaptation of the Korean version of the hip outcome score. *Knee Surg Sports Traumatol Arthrosc*. 2015 Nov;23(11):3426–3431.
- Naal FD, Impellizzeri FM, Miozzari HH, Mannion AF, Leunig M. The German Hip Outcome Score: validation in patients undergoing surgical treatment for femoroacetabular impingement. Arthroscopy. 2011;27(3):339–345.
- 22. Seijas R, Sallent A, Ruiz-İbán MA, et al. Validation of the Spanish version of the hip outcome score: a multicenter study. *Health Qual Life Outcomes*. 2014 May;13(12):70–75.