

# PSIHOMETRIČNE LASTNOSTI SLOVENSKEGA PREVODA LESTVICE DNEVNIH OPRAVIL ZNOTRAJ OCENE IZIDA KOLKA PRI STAREJŠIH OSEBAH Z OKVARO KOLKA

## PSYCHOMETRIC PROPERTIES OF THE ACTIVITIES OF DAILY LIVING SCALE OF THE SLOVENIAN VERSION OF THE HIP OUTCOME SCORE IN ELDERLY PERSONS WITH HIP DISORDERS

Petra Josipović<sup>1,2</sup>, mag. fiziot., doc. dr. Metka Moharić<sup>1,3</sup>, dr. med., Dea Salamon<sup>1,4</sup>, mag. fiziot.,  
prof. dr. Gaj Vidmar<sup>1,3,5</sup>, univ. dipl. psih.

<sup>1</sup>Univerza v Ljubljani, Medicinska Fakulteta, Ljubljana, Slovenija

<sup>2</sup>Dnevni center za rehabilitacijo Veruda, Pula, Hrvatska

<sup>3</sup>Univerzitetni rehabilitacijski institut Republike Slovenije Soča, Ljubljana, Slovenija

<sup>4</sup>Univerza na Primorskem, Fakulteta za vede o zdravju, Izola, Slovenija

<sup>5</sup>Univerza na Primorskem, FAMNIT, Koper, Slovenija

### Abstract

#### Background:

The Hip Outcome Score (HOS) is a frequently used self-administered clinical assessment tool for degenerative hip diseases. The aim of our study was to provide a reliable and valid Slovenian version of the Activities of Daily Living (ADL) scale of the HOS for use in the elderly population.

#### Methods:

The HOS ADL scale was translated and minimally adapted to Slovenian. Its metric characteristics were tested in 42 elderly patients (9 men and 33 women, age 63-99 years, median 84 years), to whom it was administered twice with a ten-day interval. Reliability, responsiveness, construct validity and convergent validity of the scale were assessed.

#### Results:

The estimated internal consistency was excellent (Cronbach alpha 0.95). The test-retest reliability was also nearly perfect (intraclass correlation 0.98). Correlations of the HOS ADL

### Izvleček

#### Uvod:

Ocena izida kolka (angl. Hip Outcome Score, HOS) je pogosto uporabljano samoocenjevalno klinično mersko orodje za degenerativne bolezni kolka. Z raziskavo smo želeli preveriti, ali je slovenska oblika lestvice dnevnih opravil (angl. Activities of Daily Living, ADL) znotraj HOS zanesljiva in veljavna za uporabo pri starejših osebah.

#### Metode:

Lestvico ADL HOS smo prevedli v slovenščino in jo nekoliko prilagodili. Njene merske značilnosti smo preverili na 42 starejših pacientih (9 moških in 33 ženskah, starosti 63-99 let, mediana 84 let), ki so jo izpolnili dvakrat z desetdnevnim presledkom. Ovrednotili smo zanesljivost, odzivnost ter konstruktno in konvergentno veljavnost lestvice.

#### Rezultati:

Notranja skladnost ocen izidov HOS ADL je bila odlična (koefficient alfa 0,95). Zanesljivost ocen ob ponovnem testiranju je

scale with the WOMAC scale (-0.80) and VAS pain assessment (-0.57) were high and statistically significant. Among SF-36 questionnaire's quality-of-life domains, we observed the lowest correlation with Social Functioning (0.30) and the highest correlation with General Health (0.66). The estimated minimum detectable change for HOS ADL was 12 points. No floor or ceiling effects were observed.

### Conclusion:

The Slovenian version of the HOS ADL demonstrated a high level of reliability and validity in the elderly population, so we recommend it for clinical use.

### Key words:

hip; elderly; Slovenia; Hip Outcome Score; reliability; validity

*bila prav tako skoraj popolna (intraklasni korelacijski koeficient 0,98). Korelacija lestvice HOS ADL z lestvico WOMAC (-0,80) in oceno bolečine na vidni analogni lestvici (-0,57) je bila visoka in statistično značilna. Med področji kakovosti življenja, ki jih meri vprašalnik SF-36, je bilo s HOS ADL najmanj povezano socialno funkcioniranje (0,30), najbolj pa splošno zdravje (0,66). Ocenjena najmanjša zaznavna sprememba na lestvici HOS ADL je znašala 12 točk. Učinka tal ali stropa nismo opazili.*

### Zaključek:

*Slovenska različica lestvice HOS ADL je visoko zanesljiva in veljavna pri starejših osebah, zato jo priporočamo za klinično uporabo.*

### Ključne besede:

*kolka; starejši; Slovenija; Ocena izida kolka; zanesljivost; veljavnost*

## INTRODUCTION

Patient-reported outcome measures (PROMs) are increasingly being used to evaluate clinical outcomes in orthopaedics, physical medicine and rehabilitation (1). PROMs are measurement tools that clinicians use to provide information on aspects of patient health status that are relevant to their quality of life, including symptoms, functionality, and physical, mental and social health (2). PROMs are important to improve patient-provider communication, patient involvement in decision-making and to better comprehend whether health care services and interventions enhance patients' health status and quality of life (2).

The growing popularity of therapeutic hip interventions continues to drive outcome-related research, primarily due to greater costs and risks associated with surgery in elderly patients with hip disorders (3-6). Hip osteoarthritis (OA) is the most common of the musculoskeletal disorders affecting the elderly (7) which can contribute to inactivity with ageing, and consequently to pain and reduced function, thus limiting the ability to perform simple activities of daily living (ADL) and ultimately impairing quality of life (8). There is a significant economic burden associated with hip disorders, largely due to the effects of disability associated with musculoskeletal disorders, comorbid diseases and cost of treatment (9).

A number of validated PROMs are currently in use for elderly with hip disorders (3); among them, the most prominent are the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) (10), the Oxford Hip Score (OHS) (11), the modified Harris Hip Score (mHHS) (12), the Hip Disability and Osteoarthritis Outcome Score (HOOS) (13), and the Hip Outcome Score (HOS) (14). The HOS was the most commonly reported PROM

utilised in assessment of ADL in the elderly with hip osteoarthritis according to the literature (15, 16). The HOS was developed in the United States of America in 2006 and it is focused on activities of daily living and sports in the general population. The HOS is a self-administered, widely used clinical assessment tool for patients with degenerative hip diseases that is short, comprehensible, and easy to administer and interpret (14, 16).

The HOS has been validated in individuals after arthroscopy and those with acetabular labral tears (17, 18). The HOS exhibits high observer agreement, internal consistency, test-retest reliability, construct validity and interpretability, and low measurement error (17, 18). Despite its widespread use for assessing patients with hip pathologies, an official Slovenian version has not been prepared and validated.

The aim of our study was to provide a reliable and valid Slovenian version of the HOS. We focused our effort only on the ADL scale because hip pathology is most often present and assessed among the elderly, where the most common indicators of functional limitations are changes in the ability to perform ADL (8, 19).

## METHODS

The quality of a PROM is assessed by means of several measurement properties. The main measurement properties are reliability (internal consistency and test-retest reliability), validity (content validity and construct validity) and responsiveness (20). When an assessment protocol is translated into another language and subsequently validated, a standardized methodology for cross-cultural adaptation should be followed (21). Therefore, we translated the HOS ADL scale into the Slovenian language in concordance with

other translation studies into Spanish, Korean, Portuguese, German and Turkish (22-26). Considering the psychometric information on the HOS and currently existing validation studies protocol, we chose the WOMAC, the Short Form Health Survey (SF-36) (27) and Visual Analogue Scale (VAS) (28) to be compared to the HOS ADL scale, which have been culturally adapted and validated in Slovenian language.

## Participants

The patients were regular residents of the nursing home Lucija in Portorož, where the study was performed. The inclusion criteria for the patients were: hip OA, femoral fracture, osteoporosis, avascular necrosis, hip pain, congenital dislocation of hip, hip effusion, muscle tear, oedema of femoral head, or acetabular cystic lesion. All the eligible patients were asked to read and sign an informed consent form that had been approved by the Slovenian National Medical Ethics Committee (0120-46/2019/19).

Out of the 85 elderly patients with different hip pathologies who were initially considered for inclusion, 31 did not meet the inclusion criteria. Those participants were excluded from the study due to the inability to cooperate, understand and fulfil the questionnaires or understand the Slovenian language, or being unable to participate in the study for other reasons (medical conditions, being alcohol or substance dependent, current alcohol or substance abuse, cardiac or other medical instability, immobilised, having active malignancy, or mental illness). Among the eligible patients, 12 refused to participate in the study. Finally, 42 elderly patients with different hip pathologies were enrolled into the study. The sample size was sufficient according to established recommendations (29).

## Instruments

The HOS is a self-administered instrument which evaluates the outcomes of treatment interventions for patients with hip disorders. The HOS is divided into two subscales, ADL (19 items) and Sports (9 items), totalling 28 items (15, 16). The HOS ADL subscale includes 19 items about ADL that are rated from 0 ("unable to do") to 4 ("no difficulty"). The total score is the rounded percentage of the maximum possible score (which depends on the number of answered items, whereby at least 14 items must be answered for a valid scoring), thus ranging from 0 to 100, with higher scores representing better function (16).

The WOMAC is a self-administered, disease-specific measure that contains subscales for pain, stiffness, and physical function (10). The global score, which we used in our study, is calculated as the sum of the scores for each subscale and ranges from 0 to 96 (10). The higher the score, the worse the health state.

The SF-36 comprises eight domains: physical functioning (PF), role limitations due to physical function (RP), bodily pain (BP), general health perceptions (GH), vitality (VH), social function (SF), emotional function (RE), and mental health (MH) (27). Each domain score is transformed to a scale from 0 to 100.

The VAS is regularly used in the nursing home where the participants reside for assessing pain intensity. The score is determined by measuring the distance (mm) on the 10-cm line between the "no pain" anchor and the patient's mark, thus providing a range of scores from 0-100. A higher score indicates greater pain intensity (28).

## Procedure

Translating the HOS ADL scale followed the established guidelines for validation and cross-cultural adaptation (20, 21). It comprised four stages: initial translation, back-translation, preparation of a consensus version, and publishing of the final version. Two Slovenian health professionals (a physiotherapist and an occupational therapist) fluent in English individually translated the scale. Their versions were compared and reviewed by a bilingual person in order to establish the first Slovenian translation.

Next, two native English speakers with a good command of the Slovenian language separately translated that translation back into English. Neither of the translators had access to the original HOS or was acquainted with the purpose of the study. The back-translated version was compared to the English original. In the end, all corrections were collected and the final translation of the HOS ADL scale into Slovenian was created (Appendix). The only small difference from the original is that in the Slovenian version, the term "Average" in the statement "Getting in and out of an average car", is replaced by a more appropriate term "normal" (the translated statement reads "Vstopanje in izstopanje iz običajnega avtomobila").

Upon first assessment, the participants completed all four instruments (HOS ADL, WOMAC, SF-36 and VAS). Ten days later, the participants completed the Slovenian HOS ADL scale again. One physiotherapist and one occupational therapist provided assistance with reading, writing, and explanation if requested. The data were collected between January 7 2020 and February 31, 2020.

## Data analysis

Reliability, responsiveness, construct validity, convergent validity, and ceiling and floor effects of the Slovenian version of the HOS ADL scale were assessed. IBM SPSS Statistics 26 software (IBM Corp., Armonk, NY, UAS) was used for statistical analysis. Descriptive statistics were calculated for all the variables. The level of significance was set at  $p < 0.05$ . For numeric variables, the Kolmogorov-Smirnov test was used for assessing departures from the normal distribution.

Reliability comprises consistency, repeatability, and agreement of experimental results (29). Internal consistency, which reflects the ability of the scale's items to measure the same construct, was assessed using Cronbach's alpha ( $\alpha$ ) coefficient, whereby a value of 0.70-0.95 was considered adequate (30). The intra-class correlation coefficient (ICC) was used to assess the test-retest reliability of the Slovenian HOS ADL scale; an ICC is usually considered good when it ranges between 0.6 and 0.9 (31), and

excellent when it exceeds 0.95, so we applied the  $ICC > 0.65$  criterion.

To assess responsiveness of the scale, standard error of measurement (SEM) and minimum detectable change (MDC, which refers to the smallest amount of change that is outside the measurement error) were estimated. SEM was calculated as the standard deviation of the scores multiplied by the square root of  $(1 - ICC)$ . The SEM was used to determine the MDC at the 95% limits of confidence (MDC95%) as  $SEM \times 1.96 \times \sqrt{2}$  (32).

Pearson's correlation coefficients were calculated to assess construct validity and convergent validity. The construct validity of the Slovenian HOS ADL scale was determined in relation to the WOMAC, the VAS and the Physical Component Summary (PCS) of the SF-36. The SF-36 domain scores and the Mental Component Summary (MCS) were used to assess convergent validity (22, 32, 33).

Floor and ceiling effects were assessed by calculating the proportion of the patients attaining the minimum (score of 0) and maximum (score of 100) possible score relative to the total number of patients, respectively. They were considered unacceptable if that proportion was above 30%, and absent if it was below 15% (32, 33).

## RESULTS

The median age of the participants was 84 years (range 63-99 years). The median age of the male participants was 81 (range 65-87) and the median age of the female participants was 84 (range 63-99). The most frequent diagnosis among the participants was hip OA (50 %) (Table 1).

**Table 1.** Diagnoses of hip disorders of the included patients.

**Tabela 1.** Diagnoze okvar kolka vključenih pacientov.

Diagnosis / Diagnoza	n	%
Hip osteoarthritis / osteoartritis kolka	21	50 %
Osteoporosis / osteoporoza	8	19 %
Avascular necrosis / avaskularna nekroza	6	14 %
Femoral fracture / zlom stegenice	3	7 %
Hip labral tear / raztrganina labruma kolka	2	5 %
Hip dysplasia / displazija kolka	1	2 %
Rheumatoid arthritis / revmatoidni artritis	1	2 %

The mean HOS ADL score (at first assessment) was 53.0 (median 61.8, range 1.5 - 100.0), the mean WOMAC score was 34.4 (median 34.9, range 1.0 - 97.9), and the mean VAS score was 3.4 (median 3.0, range 0 - 10). The distribution of HOS ADL and VAS scores was somewhat left-skewed (negative kurtosis, but not exceeding 1.0 in absolute value; statistically significant departure from normality with  $p = 0.006$  and  $p = 0.002$ , respectively).

The estimated internal consistency of the Slovenian HOS ADL scale was  $\alpha = 0.95$ . The estimated test-retest reliability was  $ICC = 0.98$ . Hence, the estimated SEM was 4.3, and MDC95% was 12.0.

The correlations of the HOS ADL scores with the WOMAC ( $r = -0.80, p < 0.001$ ) and VAS scores ( $r = -0.57, p < 0.001$ ) were high and statistically significant. The correlations of the HOS ADL scores with the SF-36 domains are reported in Table 2. The lowest correlation was with Social Functioning (about 0.3, on the margin of statistical significance), while the highest was with General Health (about 0.7, clearly statistically significant). The HOS ADL scores correlated more with the PCS ( $r = 0.64, p < 0.001$ ) than with the MCS of the SF-36 ( $r = 0.51, p = 0.001$ ), though both correlations were statistically significant.

**Table 2.** Correlations of the Slovenian HOS ADL scale scores with the SF-36 domain scores.

**Tabela 2.** Povezanost rezultatov slovenskega prevoda Lestvice dnevni opravil Ocene izida kolka s področji vprašalnika SF-36 o kakovosti življenja.

SF-36 domain / Področje	r	p
Physical Functioning / Telesno funkcioniranje	0.38	0.012
Physical Role / Telesna vloga	0.56	<0.001
Vitality / Vitalnost	0.51	0.001
Social Functioning / Socialno funkcioniranje	0.30	0.053
Emotional Role / Čustvena vloga	0.43	0.005
Mental Health / Duševno zdravje	0.53	<0.001
General Health / Splošno zdravje	0.66	<0.001
Bodily Pain / Telesna bolečina	0.38	0.013

No patient had the minimum possible HOS ADL score either at the initial or at the repeated assessment. One patient (2 %) obtained the maximum possible score at both assessments.

## DISCUSSION

The final version of the Slovenian HOS ADL subscale demonstrated an excellent level of reliability in terms of both internal consistency (Cronbach alpha of 0.95) and stability over time (ICC of 0.98).

The latter is similar to previous reports, where the time interval between assessments was from 10 to 15 days (22-26). If re-test intervals are very short, the participants can answer the questions simply based on their memory of the first assessment, while longer intervals can carry the risk of spontaneous improvement of the participant's condition. The participants in our study did not undergo interventions that would imply rapid changes in condition, and 50 % of them were diagnosed with hip OA. This means that their hip condition was not likely to change significantly over a period of one or two weeks period, so we chose the time interval

of 10 days based on similar studies (34). In addition, a study (35) reported no clinically or statistically significant difference between the measurements of test-retest reliability performed with a 2-day interval as compared with a 2-week interval for athletic patients with disorders in their study.

The SEM can be used to generate the MDC, which is the minimal amount of change in the score of an instrument that must occur in an individual in order to be sure that the change in score is not simply attributable to measurement error (22, 33). The estimated SEM for the HOS ADL scale was 4.3 and the MDC95 % was 12.0. The estimated MDC of the Spanish translation was 13.7 (22), and similar findings have been reported by others (23-26). This speaks in favour of adequacy of our translation.

The HOS ADL scale was highly negatively correlated with the WOMAC scale (about -0.8) and VAS pain assessment (about -0.6). The negative correlations were expected because a higher HOS ADL scale score indicates better physical functioning, comprising less pain, while a higher WOMAC score indicates more physical disability and a higher VAS score indicates more pain related to hip OA. As the HOS ADL scores increase, individuals experience less pain, improved physical function, and better overall quality of life related to their hip condition. The negative correlation of the HOS ADL scale with WOMAC and VAS pain assessment supports the notion that the HOS ADL scale effectively captures hip-related outcomes, as it aligns with expected relationships between different measures of pain and disability associated with hip OA. Similar results have been reported in Spain and Korea (22, 23).

The pattern of the correlations of the Slovenian HOS ADL with the SF-36 domains was not entirely expected. The higher correlation observed between the HOS ADL and the physical functioning domain of the SF-36 in the literature (nearly 0.8) (22) suggests that the relationship between hip function and physical functioning may vary across different populations or cultural contexts. It could indicate that cultural, social, or environmental factors specific to the Slovenian population might influence the association between hip function and physical functioning. These factors could include differences in lifestyle, access to healthcare, or specific demands on physical activities.

The higher correlation observed between the HOS ADL and the mental health domain (about 0.5) compared to the physical functioning domain (about 0.4) is also intriguing. It may suggest that the impact of hip function on mental well-being or psychological factors related to hip conditions is relatively stronger in the Slovenian population. This finding could have implications for understanding the holistic impact of hip conditions on individuals' overall health and quality of life. It might also indicate that interventions targeting mental health aspects in addition to physical functioning could be particularly relevant and effective in this population.

The lowest correlation observed between the HOS ADL and the social functioning domain (about 0.3) may indicate that hip function has a weaker association with social activities and interactions

compared to other domains of the SF-36. This finding could suggest that hip conditions may have less impact on individuals' social functioning or that social factors may play a relatively smaller role in influencing the relationship between hip function and social activities. Further investigation could help identify specific reasons for this weaker association and shed light on potential social factors that may mitigate or exacerbate the impact of hip conditions on social functioning.

The high correlation observed between the HOS ADL and the general health domain (nearly 0.7) supports the convergent validity of the HOS ADL scale. This finding suggests that the HOS ADL is capturing aspects of general health that are relevant to individuals with hip conditions. It underscores the importance of considering general health perceptions and overall well-being when evaluating the impact of hip conditions on patients' lives.

None of participants scored 0 points on the HOS ADL scale at either assessment, and only one participant scored 100 points at both assessments. Although only one participant achieved the maximum score of 100 points at both assessments, other participants have achieved scores close to 100. The observed variability in scores indicates that the scale is sensitive enough to detect a range of functional abilities and capture improvements or declines in participants' hip function. It suggests that the scale is responsive to changes over time and can capture meaningful differences in participants' outcomes, allowing for a comprehensive evaluation of their hip-related function and disability.

The limitations of this study include the fact that there was a heterogeneity of hip conditions with a high proportion of patients with hip OA (50 %), which may not be representative of the general population. In addition, only about one half of the initially considered patients were enrolled into the study, which may have caused selection bias. However, neither the relatively small sample size nor the heterogeneity of hip conditions decreased the estimated test-retest reliability of the HOS ADL scale. The scale was re-administrated by two healthcare professionals who regularly work with the included patients, which may have led to observer bias (33). Nevertheless, the study strongly supports the overall validity of the Slovenian translation of the HOS ADL scale for clinical use in the studied population.

## CONCLUSION

The metric characteristics of the Slovenian version of the HOS ADL scale applied to the elderly population turned out to be excellent. Internal consistency and test-retest reliability of the scale were nearly perfect, and no floor or ceiling effect was observed. The scale scores highly and statistically significantly correlated with the WOMAC scores and VAS assessment of pain, thus demonstrating construct validity. The HOS ADL scale was also statistically significantly (albeit moderately) correlated to the SF-36 domains, thus demonstrating convergent validity. Hence, we believe that the Slovenian version of the HOS ADL scale can be used for clinical assessment of the hip in the elderly population.

## Literature:

1. Health at a glance 2019: OECD library; 2019. Dostopno na: [https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2019\\_4dd50c09-en](https://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2019_4dd50c09-en) (citirano 15. 5. 2023).
2. Patient-Reported Outcome Measures (PROMs) for Hip and Knee Replacement Surgery: OECD Patient-Reported Indicator Surveys (PaRIS) Initiative. Ottawa: OECD: Canadian Institute for Health Information, Organisation for Economic Co-operation and Development; 2019. Dostopno na: <https://www.oecd.org/health/health-systems/OECD-PaRIS-hip-knee-data-collection-guidelines-en-web.pdf> (citirano 15. 5. 2023).
3. Tijssen M, van Cingel R, van Melick N. Patient-reported outcome questionnaires for hip arthroscopy: a systematic review of the psychometric evidence. *BMC Musculoskeletal Disord.* 2011;12:117.
4. McAlindon TE, Bannuru R, Sullivan MC, Arden NK, Berenbaum F, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. *Osteoarthritis Cartil.* 2014;22(3):363-88.
5. Fernandes L, Hagen KB, Bijlsma JW, Andreassen O, Christensen P, Conaghan PG, et al. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. *Ann Rheum Dis.* 2013;72(7):1125-35.
6. Koes BW, Van Tulder M, Lin CW, Macedo LG, McAuley J, Maher C. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. *Eur. Spine J.* 2010;19(12):2075-94.
7. Felson DT, Naimark A, Anderson J. The prevalence of knee osteoarthritis in the elderly. The framingham osteoarthritis study. *Arthritis Rheum.* 1987;30:914-8.
8. Marshall M, Watt FE, Vincent TL. Hand osteoarthritis: clinical phenotypes, molecular mechanisms and disease management. *Nat Rev Rheumatol.* 2018;14(11):641-56.
9. Lord J, Victor C, Littlejohns P. Economic evaluation of a primary care-based education programme for patients with osteoarthritis of the knee. *Health Technol Assess.* 1999;3:1-55.
10. Bellamy N, Buchanan W, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with OA of the hip or knee. *J Rheumatol.* 1988;15: 1833-40.
11. Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. *J Bone Joint Surg. Br.* 1996;78(2):185-90.
12. 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):737-55.
13. Nilsson AK, Lohmander LS, Klässbo M, Roos EM. Hip disability and osteoarthritis outcome score (HOOS)—validity and responsiveness in total hip replacement. *BMC Musculoskeletal Disord.* 2003;4(1):1-8.
14. Martin RL, Philippon MJ. Evidence of validity for the hip outcome score in hip arthroscopy. *Arthrosc.* 2007;23(8):822-6.
15. Harris M, McDonough CM, Leunig M, Lee CB, Callaghan JJ, Roos EM. Clinical outcomes assessment in clinical trials to assess treatment of femoroacetabular impingement: use of patient-reported outcome measures. *J Am Acad Orthop Surg.* 2013;21(1):39.
16. Christensen CP, Althausen PL, Mittleman MA, Lee JA, McCarthy JC. The nonarthritic hip score: reliable and validated. *Clin Orthop Relat Res.* 2003;406(1):75-83.
17. Martin RL, Philippon MJ. Evidence of reliability and responsiveness for the hip outcome score. *Arthrosc.* 2008;24(6):676-82.
18. Martin RL, Philippon MJ. Evidence of validity for the hip outcome score in hip arthroscopy. *Arthrosc.* 2007;23(8):822-6.
19. Clynes MA, Jameson KA, Edwards MH, Cooper C, Dennison EM. Impact of osteoarthritis on activities of daily living: does joint site matter? *Aging Clin Exp Res.* 2019;31:1049-56.
20. Mokkink LB, Terwee CB, Knol DL. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol.* 2010;10:22.
21. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol.* 1993;46(12):1417-32.
22. Seijas R, Sallent A, Ruiz-Ibán MA, Ares O, Marín-Peña O, Cuéllar R, et al. Validation of the Spanish version of the hip outcome score: a multicenter study. *Health Qual Life Outcomes.* 2014;12(1):70.
23. 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;23(11):3426-31.
24. Costa RMP, Cardinot TM, Mathias LNCDC. Validation of the Brazilian version of the hip outcome score (HOS) questionnaire. *Adv Rheumatol.* 2018;58:4.
25. Naal FD, Impellizzeri FM, Miozzari HH, Mannion AF, Leunig M. The German Hip Outcome Score: validation in patients undergoing surgical treatment for femoroacetabular impingement. *Arthrosc.* 2011;27(3):339-45.
26. Polat G, Çelik D, Çil H, Erdil M, Aşık M. Evidence for reliability, validity and responsiveness of Turkish version of Hip Outcome Score. *Acta Orthop Traumatol Turc.* 2017;51(4):319-24.
27. Marn-Vukadinović D, Jamnik H. Validation of the short form-36 health survey supported with isokinetic strength testing after sport knee injury. *J Sport Rehabil.* 2011;20(3):261-76.
28. Alghadir AH, Anwer S, Iqbal A, Iqbal ZA. Test-retest reliability, validity, and minimum detectable change of visual analog, numerical rating, and verbal rating scales for measurement of osteoarthritic knee pain. *J Pain Res.* 2018;11:851.
29. Donner A, Eliasziw M. Sample size requirements for reliability studies. *Stat Med.* 1987;6(4): 441-448.
30. Terwee CB, Bot SD, de Boer MR. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42.
31. Polit DF. Assessing measurement in health: beyond reliability and validity. *Int J Adv Nurs Stud.* 2015;52(11):1746-53.
32. Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med.* 1998; 26(4):217-38.
33. Josipović P, Moharič M, Salamon D. Translation, cross-cultural adaptation and validation of the Slovenian version of Harris Hip Score. *Health Qual Life Outcomes.* 2020;18(1):1-8.
34. Khaja A, Al-Mutairi O, Al-Kudair, Al-Samhan A. Translation and cross-cultural adaptation of the Harris Hip Score into Arabic. *Glob J Med Res.* 2020;20:19-26.
35. Marx RG, Menzes A, Horovitz L, Jones EC, Warren RF. A comparison of two-time intervals for test-retest reliability of health status instruments. *J Clin Epidemiol.* 2003;56(8):730-5.

## Appendix: Slovenian translation of the Hip Outcome Score Activities of Daily Living scale

### Lestvica izida kolka – Lestvica dnevnih aktivnosti

Prosimo, odgovorite na vsako vprašanje z odgovorom, ki najbolj ustreza opisu vašega stanja v preteklem tednu. Če so vaše dnevne aktivnosti omejene zaradi drugega razloga, ki ni posledica obolelega kolka, označite "Ni odgovora [x]".

Koliko težav imate pri izvajanju naslednjih dejavnosti zaradi obolelega kolka?

	Ni težav (4)	Rahle težave (3)	Zmerne težave (2)	Resne težave (1)	Nezmožen izvedbe (0)	Ni odgovora (x)
1. 15 minut stoje						
2. Vstopanje in izstopanje iz običajnega avtomobila						
3. Hoja po strmini navzgor						
4. Hoja po strmini navzdol						
5. Vzpenjanje za eno nadstropje stopnic						
6. Spuščanje za eno nadstropje stopnic						
7. Stopanje na robnik ali z njega						
8. Globoki počep						
9. Vstopanje in izstopanje iz kadi						
10. Začenjanje hoje						
11. Hoja okrog 10 minut						
12. Hoja 15 minut ali več						
13. Obračanje/vrtenje na bolni nogi						
14. Obračanje v postelji						
15. Lahko do zmerno težko delo (stoja/hoja)						
16. Težko delo (potiskanje/vlečenje, plezanje, nošenje bremen)						
17. Rekreativna						

Kako bi ocenili vaše delovanje pri običajnih dnevnih aktivnostih med 0 in 100, če 100 pomeni izvedbo aktivnosti brez težav, 0 pa pomeni, da izvedba dnevnih aktivnosti ni možna? \_\_\_\_\_

Ali imate zaradi svojega kolka težave z naslednjima dejavnostnima?

	Ni težav	Rahle težave	Zmerne težave	Resne težave	Nezmožen izvedbe	Ni odgovora
Obuvanje nogavic in čevljev [brez točkovanja]						
15 minut sedenja [brez točkovanja]						