

# An update on joint-specific outcome measures in total hip replacement

Łukasz Pulik<sup>1</sup>, Krzysztof Romaniuk<sup>1</sup>, Kaja Jaśkiewicz<sup>2</sup>, Paweł Wojtyński<sup>1</sup>, Paweł Łęgosz<sup>1</sup>, Paweł Małydk<sup>1</sup>

<sup>1</sup>Department and Clinic of Orthopedics and Traumatology, 1<sup>st</sup> Faculty of Medicine, Medical University of Warsaw, Poland

<sup>2</sup>Department and Clinic of Internal Medicine, Pulmonary Diseases and Allergy, Medical University of Warsaw, Poland

## Abstract

This is the latest review of joint-specific tools used to evaluate patients undergoing total hip replacement (THR) surgery, which is an effective treatment for end-stage osteoarthritis. Due to the large number and multitude of scales and their variants used, a critical assessment of the available tools is necessary. In the article, we briefly describe six different clinical tools: the Western Ontario and McMaster Universities Osteoarthritis Index, the Hip Disability and Osteoarthritis Outcome Score, the Harris Hip Score, the Oxford Hip Score, the Mayo Hip Score, and the Rheumatoid and Arthritis Outcome Score. We present the advantages and constraints of the different outcome measures, providing a helpful resource of information for clinical trials and for everyday routine evaluation.

**Key words:** arthroplasty, patient outcome assessment, clinical scales, total hip replacements, quality of life.

## Introduction

Osteoarthritis (OA) affects more than 250 million people, representing about 6% of the global population, and the prevalence is increasing due to the aging of the population and the rising number of joint injuries [1].

Currently, in Europe 5% of men and 11% of women aged  $\geq 60$  years old suffer from symptomatic hip OA. It is one of the leading causes of disability, limiting everyday activities, and a factor affecting patients' quality of life (QoL).

In Poland, over 160 total hip replacements (THR) are performed per 100,000 inhabitants per year, which ranks THR among the top ten most frequently performed surgical procedures [2]. In order to evaluate the effectiveness of various surgical techniques, rehabilitation protocols and other aspects of THR, comprehensive, and relevant outcome measures are necessary.

Through the course of history, the focus of evaluation of the treatment outcomes shifted from observer reported outcomes (OROs), which described mainly the procedural effect of the surgery, to patients' opinion on the treatment, obtained through questionnaires collectively

named "Patient Related Outcome Measures" (PROMs). This approach focuses on factors such as pain, the degree of disability, and the impact of surgery results on daily activities and quality of life [3].

The outcome measures can be divided into two other broad categories depending on the scope of health issues addressed – joint-specific, and generic health-related, measuring the overall state of well-being. In the assessment of THR results, both types are recommended for the complex and accurate evaluation of the treatment effect [4].

This review aims to focus on joint-specific outcome measures. Some of the clinical scales' forms are difficult to access, so researchers tend to use questionnaires from sources other than the original, which can lead to unintentional modifications [5].

Also, there are significant inconsistencies and lack of clarity in reporting of the outcome measures scoring methods. Similarly, missing values are addressed differently among different clinical trials. It results in uncertainty in the interpretation of study results and puts limitations on data synthesis across different trials [6].

---

## Address for correspondence:

Łukasz Pulik, Department and Clinic of Orthopedics and Traumatology, 1<sup>st</sup> Faculty of Medicine, Medical University of Warsaw, 4 Lindley St., 02-005 Warsaw, Poland, e-mail: [lukaszpulik@gmail.com](mailto:lukaszpulik@gmail.com)

Submitted: 27.01.2020; Accepted: 18.03.2020

This state-of-the-art review will describe the most common joint-specific tools used to evaluate various areas of treatment in patients undergoing THR.

## Western Ontario and McMaster Universities Osteoarthritis Index

**Background:** The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) developed by Bellamy et al. in 1982 is a widely used disease-specific questionnaire applied to measure the change in clinical status in the treatment of hip and knee OA. It is recommended by the Osteoarthritis Research Society International clinical trial guidelines for the outcome measurement in OA [7]. Its validity was proven in orthopedic outcome studies in the assessment of the effectiveness of surgery such as THR and total knee replacement (TKR), pharmacotherapy, and exercise therapy in the treatment of OA [6, 8]. It is a self-administered PROM and requires no help from a skilled physician or physiotherapist [9].

**Contents:** The WOMAC consists of 24 items divided into three domains, which are pain (5 questions), stiffness (2 questions), and physical function (17 questions) [9, 10].

**Recall period:** Forty-eight hours for the standard version. Currently, there are also 24 hours, past seven days, and past month versions available [11].

**Time to complete:** Average time to complete ranges from 3 to 6 minutes for the paper version of WOMAC, mobile 4.5–5 minutes, and for patient interview 5–7.5 minutes [12].

**Available versions:** Paper [9], emailed [13], mobile phone app [14], and touch-screen computerized [15]. WOMAC also can be completed over the telephone [16].

**Available Scoring Variants:** Available for use are versions evaluating the condition of the patient using a 5-point Likert scale [10], 100 mm horizontal visual analog scale [10, 17], and eleven-point numerical format [15]. There is also a version of WOMAC with signal items versus complete index usage [11].

**Scoring:** The maximum score for the Likert variant of the questionnaire (41% of the studies) is 96 points (worst function, stiffness, severe pain). The fewer points the patient gains from answers to the questionnaire, the better the condition of the joint or the outcome. The WOMAC is divided into three subscales: 20 points for pain, 8 for stiffness, and 68 for function. The total score in the Likert variant can be calculated by summing up all points (0–96 points), the average value (max. of 4), or the percent of maximum value. The standard procedure for missing data in the WOMAC states that the average score from the subscale substitutes the missing value. However, when the subject did not answer  $\geq 4$  of the 17

function questions, the subscale is invalid – the same for the stiffness subscale ( $\geq 2$ ) and pain ( $\geq 2$ ) [18].

**Languages:** The WOMAC is one of the most frequently used and recognized evaluation tools in lower limb OA. To date, there are 85 different language versions. The WOMAC developer also provides a Polish language version, but to our knowledge, there is no validation study [18].

**How to obtain and licensing:** The information regarding the WOMAC 3.1 file and license to use for research are available on the [www.womac.com](http://www.womac.com) website. All the questions of the Likert variant of WOMAC (version 3.0) are included in their original form in the Hip Disability and Osteoarthritis Outcome Score questionnaire [19].

**Psychometrics:** The internal consistency reliability for the Likert version of WOMAC was Cronbach's  $\alpha = 0.95$  for function, 0.83 for pain, and 0.81 for stiffness. The test-retest reliability was respectively 0.91 for function, 0.87 for pain, and 0.80 for stiffness subscales [12]. The WOMAC is a valid and reliable tool for the assessment of the severity of OA [10]. However, its ability to detect change is limited due to the overlap of pain and function items [20].

**Variants of the original questionnaire:** There are various short-form versions of the questionnaire validated (WOMAC-SF) [21]. In 2003 Whitehouse and colleagues validated reduced WOMAC function. The reduced version retained 7 of the original 17 items from the function subscale, which are: ascending stairs, rising from sitting, walking on the flat, getting in or out of a car, putting on socks, rising from bed, and sitting. Reduced WOMAC function has been shown as valid and reliable as a full version of WOMAC and presented higher responsiveness than the full WOMAC for THR and TKR patients (1.4 vs. 1.6) [22, 23].

## Hip Disability and Osteoarthritis Outcome Score

**Background:** The Hip Disability and Osteoarthritis Outcome Score (HOOS) is a joint-specific self-administered clinical scale. The HOOS includes all the WOMAC version 3.0 questions (with permission). It also addresses issues relevant to younger, more active patients, such as sport and recreation function [19].

**Contents:** HOOS has 40 items grouped into five subscales: pain, other symptoms, function in daily living, function in sports and recreation, and hip-related QoL [19].

**Recall period:** When answering the questions last week is taken into consideration.

**Time to complete:** From 10 to 15 minutes to complete [24].

**Available versions:** Paper version, automated mobile phone messaging robot version of the shortened HOOS physical function subscale (HOOS-PF), and HOOS Pain subscale [25].

**Available scoring variants:** Answers are divided into five Likert boxes.

**Scoring:** The score of each subscale is a sum of all the points obtained (0–4 points per question). The total score is the sum of all subscales transformed and expressed as a score ranging from 0 to 100 (worst to best possible outcome). Missing data are addressed with the implementation of mean values for a subscale if 50% or more items are answered. The score can be calculated with an Excel tool or manually – mean points from the subscale divided by four multiplied by 100 and then subtracted from 100. WOMAC can be calculated from HOOS. Instructions are available on the [www.koss.nu](http://www.koss.nu) website [19].

**Languages:** HOOS was initially developed in English. Currently, there are 19 languages available. Also a validated cross-cultural adapted Polish version is available [26].

**How to obtain and licensing:** On the [www.koss.nu](http://www.koss.nu) website, no license is required.

**Psychometrics:** The Cronbach's  $\alpha$  coefficient for internal consistency ranges from 0.82 to 0.98. HOOS is validated for two versions, LK1.1 and LK 2.0. HOOS has been tested against SF-36, and correlations indicating its validity were found. HOOS is valid against the Short Form 36, the Oxford Hip Score, the Lequesne Index, and the visual analog scale. The response rate varied between studies from 1.29 to 3.24. Test-retest reproducibility is characterized by a level of interclass correlation of  $> 0.78$  [24].

**Variants of the original questionnaire:** HOOS Joint Replacement (JR) is a short version of HOOS and consists of six items concerning function and pain. It provides a comparable level of psychometric performance to the complete HOOS. It is recommended by the Centers for Medicaid and Medicare Services and the American Association of Hip and Knee Surgeons for patients undergoing THR. The HOOS JR can be accessed through <https://www.hss.edu/hoos-jr-kooos-jr-outcomes-surveys.asp> website [24]. There is also a twelve-item version of the questionnaire, HOOS-12, which is a valid and reliable alternative to the 40-item HOOS in THR [24]. Also, HOOS-PF – the aforementioned shortened version of the HOOS physical function subscale – was developed, and it consists of 5 items [27].

## Harris Hip Score

**Background:** The Harris Hip Score (HHS) is a joint-specific, ORO measure developed in 1969 as a tool to evaluate the results of hip surgery, and it consists of two sections:

questions and a physical examination, including a range of motion and deformity items, which differentiates HHS from other presented clinical scales. Currently, HHS is the most widely used hip rating scale for THR patients [28].

**Contents:** The 11 items are grouped into three sections: pain (1 item), function and everyday activities (7 items), physical examination (3 items). The range of motion is measured for flexion, abduction, adduction, and external rotation [5].

**Recall period:** Unspecified [5].

**Time to complete:** It takes about 5 minutes to complete [5].

**Available versions:** Paper HHS and telephone-call version of modified HHS tested in patients after THR [29].

**Available scoring variants:** Likert-type boxes, but the number of boxes varies depending on the question [5].

**Scoring:** The score has a maximum of 100 points (best possible outcome) with a maximum of 44 points for pain, 47 for function, 4 points for absence of deformity, and 5 points for a range of motion. The highest score of 100 points indicates the best function and no pain. It should be noted, during statistical analysis, that the scoring is not continuous but rather gradual, and non-parametric tests should be used [5].

**Languages:** HHS was initially developed in English. Recently there were validation studies in Turkish [30] and Italian language [31].

**How to obtain and licensing:** The [www.orthopaedicscore.com](http://www.orthopaedicscore.com) site. No license required.

**Psychometrics:** HHS has been validated for the assessment of outcomes of the THR, OA, and femoral neck fractures [5]. The HHS validity has been verified by directly comparing HHS, the WOMAC, and the Short Form 36 (SF-36). Cronbach's  $\alpha$  coefficient for internal consistency was assessed as high [32]. The test-retest reliability of the pain subscale was  $r = 0.93$  and  $r = 0.98$ , respectively, and the function subscale  $r = 0.95$  and  $r = 0.93$ , respectively, also with good internal correlations (0.74–1.0) [33]. The response rate for the HHS was higher (1.70) than for SF-36 subscales [34].

**Variants of the original questionnaire:** In the Modified Harris Hip Score (mHHS) the original range of motion and deformity were excluded. The remaining seven items were grouped into three categories and scored using Likert boxes. The maximum score (100) indicates the best function as the original version. The mHHS can be self-reported and serve as a PROM and present similar statistical properties as the original ORO HHS [35].

## Oxford Hip Score

**Background:** The Oxford Hip Score (OHS) is a short 12-item survey, and it was proposed by Dawson et al. in 1996

[36]. It is widely used, joint-specific PROMs applied to evaluate the clinical outcomes of THR, and its validity has been proven in prospective studies for THR and TKR [36], pharmacological treatment [37] and rehabilitation [38].

**Contents:** OHS consists of 12 items concerning pain, physical function, gait, self-care, and use of a car. There are no official subscales, but five items include questions regarding pain complaints, and seven items address the function.

**Recall period:** The last four weeks are taken into consideration [39].

**Time to complete:** It takes about 5 minutes to complete [39].

**Available versions:** Paper and telephone administered version [39].

**Available scoring variants:** Five Likert boxes, with an indication of the limb affected [36].

**Scoring:** The original questionnaire's score ranged from 12 to 60. A higher score indicates worse disability. However, many surgeons modified the scoring, which leads to the creation of updated scoring with values per item ranging from 0 to 4, with 4 indicating the best state (48 points) [39]. Scores can also be converted by subtracting the score from 60 [40]. If one or two questions remain blank, the mean score is incorporated. The score is invalid when there is missing data in more than two items [36].

**Languages:** Apart from the original English, Dutch [41], Swedish [42], and other versions are available. The authors suggest the use of adaptations performed by the Haverkamp et al. method [41]. Unofficial versions from many languages are available for use and are listed on the official website of the Oxford University. There is no official Polish version.

**How to obtain and licensing:** Questionnaire and instruction on obtaining the license are available at <https://innovation.ox.ac.uk/outcome-measures/oxford-hip-score-ohs/>.

**Psychometrics:** Cronbach's measure of internal consistency was 0.84 preoperatively and 0.89 postoperatively. The OHS was validated against the Charnley score and SF-36. The OHS was more sensitive to change than SF-36 [36]. It is characterized by good test-retest reliability interclass correlation (ICC) in hip OA (ICC > 0.80) and THR (ICC > 0.70). The responsiveness was assessed as good in OA (standardized response mean, SRM = 1.12) [41, 43].

**Variants of the original questionnaire:** No other variants.

## Mayo Hip Score

**Background:** The Mayo Hip Score (MHS) devised in 1985 by Kavanagh and Fitzgerald consists of a PROM

questionnaire and radiographic evaluation developed specifically for measuring clinical outcomes of revision of total hip arthroplasty [44]. Thirty-one years later, it has been used to assess the outcomes of primary total hip arthroplasty [45].

**Contents:** Clinical Parts: pain, function (distance walked, walking aids), mobility and muscle power (getting in and out of the car, foot care, limp, climbing stairs). The MHS clinical component is similar to HHS but does not include a deformity assessment and joint range of motion [44].

**Recall period:** Unspecified by the authors [44].

**Time to complete:** The exact values of the time required to complete the scale were not measured. However, the clinical part takes 2 to 5 minutes to complete. The time required for radiological assessment has not been measured either [44].

**Available versions:** Paper version.

**Available scoring variants:** Multiple choice questionnaire with a total of 100 points in combined clinical and radiological scale.

**Scoring:** Clinical (80 points) and radiological (20 points). A higher score indicates better results. The clinical part consists of seven items with a weighted amount of points per item. It is divided into pain (40 points), function (20 points), and mobility and muscle power (20 points). Points for roentgenographic data are divided into the assessment of the condition of the acetabulum (10 points) and femur (10 points). However, scoring is only possible for cemented components [44].

**Languages:** The Mayo Hip Score was initially written in American English [44]. To our knowledge, no translations are available.

**How to obtain and licensing:** No license required. The PROMs can be found in the open-source article in Clinical Orthopedics and Related Research journal: [https://journals.lww.com/clinorthop/Abstract/1985/03000/Clinical\\_and\\_Roentgenographic\\_Assessment\\_of\\_Total.17.aspx](https://journals.lww.com/clinorthop/Abstract/1985/03000/Clinical_and_Roentgenographic_Assessment_of_Total.17.aspx).

**Psychometrics:** Mayo hip score is a valid measurement tool (against HHS) in the primary THR and also can successfully predict the risk of revision surgery. The responsiveness was assessed after 2 years (SRM = 2.61) and 5 years (SRM = 2.42) [45].

**Variants of the original questionnaire:** No other variants, besides independent use of the clinical part of the score as a PROM.

## Rheumatoid and Arthritis Outcome Score

**Background:** The Rheumatoid and Arthritis Outcome Score (RAOS) is a PROM developed to measure the severity of chronic lower limb joint issues such as rheumatoid arthritis, spondyloarthropathies, psoriatic arthri-

tis, etc. The RAOS is an adaptation of the Knee injury and Osteoarthritis Outcome Score (KOOS). By changing the word “knee” in KOOS to “leg” or “hip, knee and foot” in all of the items, a new clinical scale was devised. The RAOS validation clinical trial for patients with hip rheumatic diseases undergoing THR and TKR is ongoing according to the authors’ information [46].

**Contents:** RAOS has 42 items, and it can be divided into subscales: Pain (9), Symptoms (7), Activities of Daily Living (ADL) (17), Sports and Recreation (5) and QoL (4).

**Recall period:** When answering the questionnaire, last week is taken into consideration.

**Time to complete:** It takes about 10 minutes to complete.

**Available versions:** Only the paper version is currently available [46].

**Available scoring variants:** Answers are put into five Likert-type boxes [46].

**Scoring:** When completing the questionnaire, the maximal score indicates the worst condition. To obtain a normalized score, sum up the total score of each subscale and divide by the possible maximum score for the subscale. It includes WOMAC in its composition. The WOMAC 3.0 is included in RAOS in its full form and can be calculated from its score [46]. The scoring instructions and calculation tools (Excel) are available at [www.koos.nu](http://www.koos.nu).

**Languages:** RAOS is available in several languages, including English, Swedish, Turkish, French and Polish.

**How to obtain and licensing:** The questionnaire is available at [www.koos.nu](http://www.koos.nu). No license is required to use this questionnaire.

**Psychometrics:** The RAOS is characterized by good test-retest repeatability with Cronbach’s  $\alpha$  of 0.78–0.95. Intraclass correlations for RAOS is ICC = 2.1. The validity has been successfully tested against SF-36 the Stanford Health Assessment Questionnaire (HAQ) and the Arthritis Impact Measurement Scale (AIMS2). The responsiveness was presented as effect sizes for each of the subscales (pain 0.40, symptoms 0.41, ADL 0.44, Sports and Recreation 0.42 and QoL 0.30) [46].

**Variants of the original questionnaire:** No other variants available.

## Discussion

When trying to apply the scale, a scientist might encounter several barriers. The existence of such a large number of scales implies that it is difficult to compare both individual results and entire studies. The Oxford Hip Score and HOOS scales are used more frequently, but the Harris Hip Score (HHS) is still the most widely used scale for research, particularly randomized controlled trials, and most studies use multiple scales [47].

Although there is a growing body of activity leading to the collection of data on THR results reported by both physicians (OROs) and patients (PROMs), it remains unclear how to read these results in daily clinical practice [48].

In order to solve this problem, a cohort retrospective study was conducted in 2015 by Berliner et al. to investigate the relationship between preoperative PROM results and total hip alloplasty. The purpose was to establish a scale threshold that could predict a clinically significant improvement in the functional performance of the hip joint after THR in patients qualified for surgery. The threshold for the HOOS scale was 51.0, while for the PCS scale it was 32.5. The result of this study may support the clinician’s decision to postpone surgery if the expected probability of improvement after surgery is low, even if the symptoms of hip arthritis are severe [49]. Further research is needed to find thresholds of other scales.

For the ongoing evaluation of clinical data and to determine which factors are crucial for the patient, an international working group has been established under the OMERACT program (Outcomes 3 Measurement in Rheumatology). In 2016, a set of priority points covered by the group included evaluation of function, pain, satisfaction, revision occurrence, adverse events, and patient death [50].

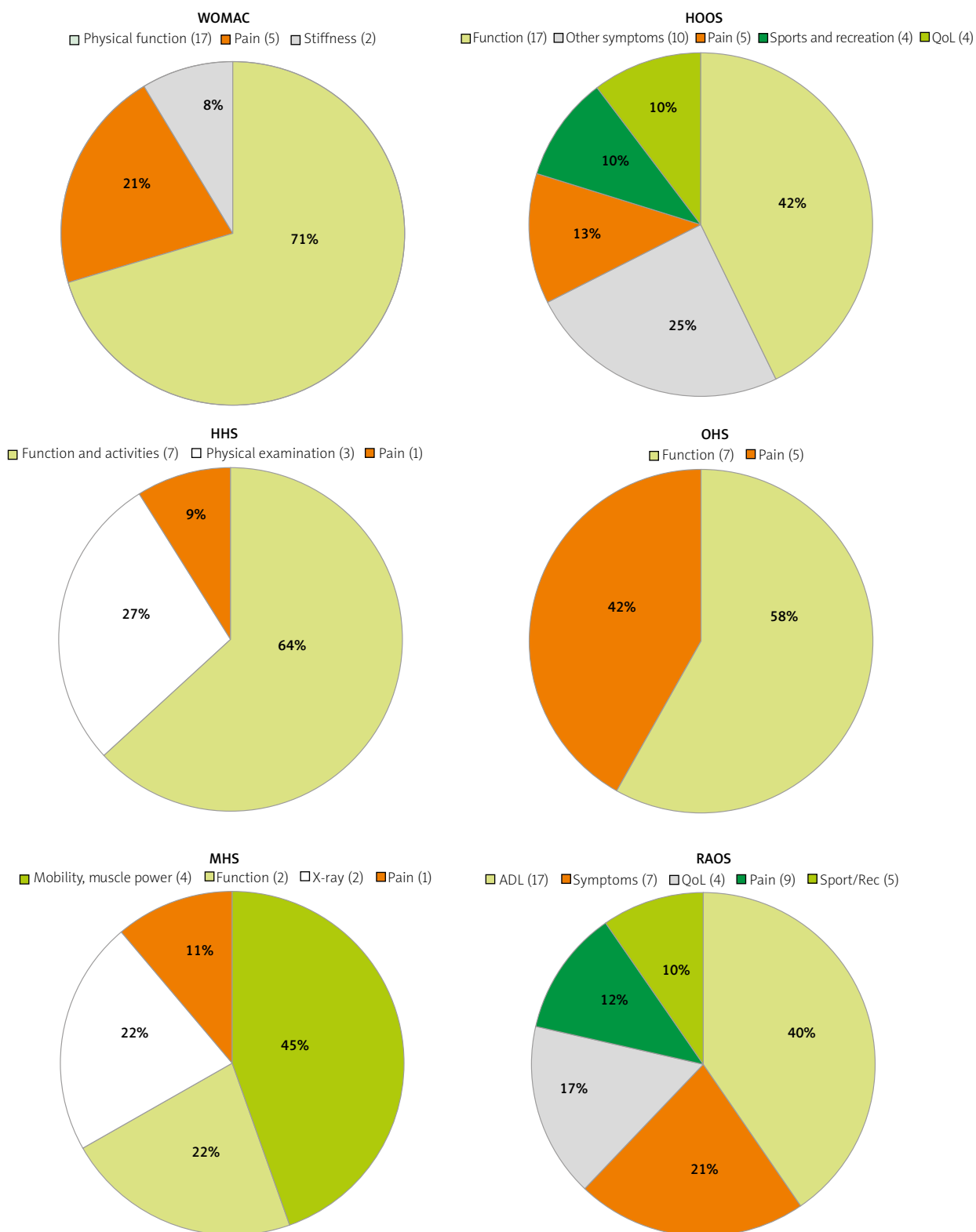
Therefore, as parameters such as pain, joint function and mobility do not fully reflect the patient’s situation, the role of patient evaluation in the evaluation of surgery results has recently been highlighted [51]. In Figure 1 items ( $n$ ) of the joint-specific outcome measures in total hip replacement are presented.

In order to increase the use of scales in the pre- and postoperative evaluation of a patient undergoing total hip alloplasty, a comprehensive strategy should be introduced to support clinicians and researchers. The share of PROMs is significantly increased by the introduction of a form on the website, which allows it to be completed at such a time as the patient needs, and allows one to monitor patients living far away or in a severe condition to come to the office only for a follow-up visit [52].

However, some patient groups may be unable to complete the form. Such groups include the elderly, other nationalities, patients with a history of three or more orthopedic procedures and undergoing revision surgery [3].

Although modern electronic devices such as telephone, e-mail or website are on the rise, contact via traditional mail remains the primary tool for long-distance postoperative contact due to the age profile of patients undergoing hip alloplasty [53].

Many factors contribute to achieving satisfactory results in total hip alloplasty. A useful tool to examine the most significant factors is outcome measures. More extensive use of these scales will involve finding patients



ADL – Activities of Daily Living, HHS – Harris Hip Score [28], HOOS – Hip Disability and Osteoarthritis Outcome Score [19], MHS – Mayo Hip Score [44], OHS – Oxford Hip Score [36], RAOS – Rheumatoid and Arthritis Outcome Score, Sport/Rec – sports and recreation subscale, QoL – Quality of Life [46], WOMAC – Western Ontario and McMaster Universities Osteoarthritis Index [9], X-ray – radiologic evaluation in aspect of loosening.

Fig. 1. Items (n) of the joint-specific outcome measures in total hip replacement.

who can achieve clinically significant improvement after surgery.

## Conclusions

In clinical practice, accurate and appropriate tools are necessary to evaluate the patient's present state of health and its changes over time. There are many different measures used for the assessment of patients undergoing THR.

In this review, we present complex and detailed information about the most frequently applied questionnaires, together with its origin, scoring methods, and psychometric properties such as validity, internal consistency, and test-retest reliability.

We also highlight the importance of standardization of the versions and scoring methods used for improved clinical applicability and comparability between different clinical trials. The advantages and limitations of the individual outcome measures are briefly explained, offering a useful source of knowledge for researchers and an everyday routine assessment of THR results.

*The authors declare no conflict of interest.*

## References

- Hunter DJ, Bierma-Zeinstra S. Osteoarthritis. *Lancet* 2019; 393: 1745-1759, DOI: 10.1016/S0140-6736(19)30417-9.
- EUROSTAT. Surgical operations and procedures performed in hospitals by ICD-9-CM. Retrieved from <https://ec.europa.eu/eurostat/data/database>, 2020.
- Schamber EM, Takemoto SK, Chenok KE, Bozic KJ. Barriers to completion of Patient Reported Outcome Measures. *J Arthroplasty* 2013; 28: 1449-1453, DOI: 10.1016/j.arth.2013.06.025.
- Rolfson O, Wissig S, van Maasackers L, et al. Defining an International Standard Set of Outcome Measures for Patients With Hip or Knee Osteoarthritis: Consensus of the International Consortium for Health Outcomes Measurement Hip and Knee Osteoarthritis Working Group. *Arthritis Care Res (Hoboken)* 2016; 68: 1631-1639, DOI: 10.1002/acr.22868.
- Banaszkiewicz PA. 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. In: *Classic Papers in Orthopaedics*, eds. Banaszkiewicz PA, Kader DF. Springer London, London 2014: 13-17, DOI: 10.1007/978-1-4471-5451-8.
- Copsey B, Thompson JY, Vadher K, et al. Problems persist in reporting of methods and results for the WOMAC measure in hip and knee osteoarthritis trials. *Qual Life Res* 2019; 28: 335-343, DOI: 10.1007/s11136-018-1978-1.
- Bellamy N. WOMAC: a 20-year experiential review of a patient-centered self-reported health status questionnaire. *J Rheumatol* 2002; 29: 2473-2476.
- Łęgosz P, Sarzyńska S, Pulik Ł, et al. Heterotopic ossification and clinical results after total hip arthroplasty using the anterior minimally invasive and anterolateral approaches. *Arch Med Sci* 2020; 16: 613-620, DOI: 10.5114/aoms.2018.78653.
- Bellamy N, Buchanan WW, Goldsmith CH, et al. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol* 1988; 15: 1833-1840.
- McConnell S, Kolopack P, Davis AM. The Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC): a review of its utility and measurement properties. *Arthritis Rheum* 2001; 45: 453-461, DOI: 10.1002/1529-0131(200110)45:5<453::aid-art365>3.0.co;2-w.
- 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): 148-153.
- Gandek B. Measurement properties of the Western Ontario and McMaster Universities Osteoarthritis Index: a systematic review. *Arthritis Care Res (Hoboken)* 2015; 67: 216-229, DOI: 10.1002/acr.22415.
- Quintana JM, Escobar A, Bilbao A, et al. Responsiveness and clinically important differences for the WOMAC and SF-36 after hip joint replacement. *Osteoarthritis Cartilage* 2005; 13: 1076-1083, DOI: 10.1016/j.joca.2005.06.012
- Bellamy N, Wilson C, Hendrikz J, et al. Osteoarthritis Index delivered by mobile phone (m-WOMAC) is valid, reliable, and responsive. *J Clin Epidemiol* 2011; 64: 182-190, DOI: 10.1016/j.jclinepi.2010.03.013.
- Theiler R, Spielberger J, Bischoff HA, et al. Clinical evaluation of the WOMAC 3.0 OA Index in numeric rating scale format using a computerized touch screen version. *Osteoarthritis Cartilage* 2002; 10: 479-481, DOI: 10.1053/joca.2002.0807.
- Lopez Alonso SR, Martínez Sánchez CM, Romero Cañadillas AB, et al. [Metric properties of WOMAC questionnaires-original and reduced versions-to measure symptoms and Physical Functional Disability]. *Aten Primaria* 2009; 41: 613-620, DOI: 10.1016/j.aprim.2009.02.005.
- Kersten P, White PJ, Tennant A. The visual analogue WOMAC 3.0 scale--internal validity and responsiveness of the VAS version. *BMC Musculoskelet Disord* 2010; 11: 80, DOI: 10.1186/1471-2474-11-80.
- Ghomrawi HM, Mandl LA, Rutledge J, et al. Is there a role for expectation maximization imputation in addressing missing data in research using WOMAC questionnaire? Comparison to the standard mean approach and a tutorial. *BMC Musculoskelet Disord* 2011; 12: 109, DOI: 10.1186/1471-2474-12-109.
- Nilsdotter AK, Lohmander LS, Klässbo M, et al. Hip disability and osteoarthritis outcome score (HOOS) – validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord* 2003; 4: 10, DOI: 10.1186/1471-2474-4-10.
- Stratford PW, Kennedy DM. Does parallel item content on WOMAC's pain and function subscales limit its ability to detect change in functional status? *BMC Musculoskelet Disord* 2004; 5: 17, DOI: 10.1186/1471-2474-5-17.
- Bilbao A, Quintana JM, Escobar A, et al. Validation of a proposed WOMAC short form for patients with hip osteoarthritis. Health and quality of life outcomes 2011; 9: 75, DOI: 10.1186/1477-7525-9-75 .

22. Whitehouse SL, Lingard EA, Katz JN, Learmonth ID. Development and testing of a reduced WOMAC function scale. *J Bone Joint Surg Br* 2003; 85: 706-711.
23. Whitehouse SL, Crawford RW, Learmonth ID. Validation for the reduced Western Ontario and McMaster Universities Osteoarthritis Index function scale. *J Orthop Surg (Hong Kong)* 2008; 16: 50-53, DOI: 10.1177/230949900801600113.
24. Nilsson A, Bremander A. Measures of hip function and symptoms: Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), and American Academy of Orthopedic Surgeons (AAOS) Hip and Knee Questionnaire. *Arthritis Care Res* 2011; 63: 200-207, DOI: 10.1002/acr.20549.
25. Scott EJ, Anthony CA, Rooney P, et al. Mobile Phone Administration of Hip-Specific Patient-Reported Outcome Instruments Correlates Highly With In-office Administration. *J Am Acad Orthop Surg* 2020; 28: e41-e46, DOI: 10.5435/JAAOS-D-18-00708.
26. Glinkowski W, Żukowska A, Dymitrowicz M et al. Translation, Cross-Cultural Adaptation, and Psychometric Properties of the Polish Version of the Hip Disability and Osteoarthritis Outcome Score (HOOS). *Medicina (Kaunas)* 2019; 55: 614, DOI: 10.3390/medicina55100614.
27. Davis AM, Perruccio AV, Canizares M, et al. The development of a short measure of physical function for hip OA HOOS-Physical Function Shortform (HOOS-PS): an OARSI/OMERACT initiative. *Osteoarthritis Cartilage* 2008; 16: 551-559, DOI: 10.1016/j.joca.2007.12.016.
28. 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: 737-755.
29. Sharma S, Shah R, Draviraj KP, Bhamra MS. Use of telephone interviews to follow up patients after total hip replacement. *J Telemed Telecare* 2005; 11: 211-214, DOI: 10.1258/1357633054068883.
30. Celik D, Can C, Aslan Y et al. Translation, cross-cultural adaptation, and validation of the Turkish version of the Harris Hip Score. *Hip Int* 2014; 24: 473-479, DOI: 10.5301/hipint.5000146.
31. Dettoni F, Pellegrino P, La Russa MR, et al. Validation and cross cultural adaptation of the Italian version of the Harris Hip Score. *Hip Int* 2015; 25: 91-97, DOI: 10.5301/hipint.5000184.
32. Nilsson A, Bremander A. Measures of hip function and symptoms: Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), and American Academy of Orthopedic Surgeons (AAOS) Hip and Knee Questionnaire. *Arthritis Care Res* 2011; 63: 200-207, DOI: 10.1002/acr.20549.
33. Söderman P, Malchau H. Is the Harris hip score system useful to study the outcome of total hip replacement? *Clin Orthop Relat Res* 2001; (384): 189-197, DOI: 10.1097/00003086-200103000-00022.
34. Hoeksma HL, Van Den Ende CH, Runday HK, et al. Comparison of the responsiveness of the Harris Hip Score with generic measures for hip function in osteoarthritis of the hip. *Ann Rheum Dis* 2003; 62: 935-938, DOI: 10.1136/ard.62.10.935.
35. Mahomed NN, Arndt DC, McGrory BJ, Harris WH. The Harris hip score: comparison of patient self-report with surgeon assessment. *J Arthroplasty* 2001; 16: 575-580, DOI: 10.1054/arth.2001.23716.
36. 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: 185-190.
37. Clarke S, Lock V, Duddy J, et al. Intra-articular hylan G-F 20 (Synvisc) in the management of patellofemoral osteoarthritis of the knee (POAK). *Knee* 2005; 12: 57-62, DOI: 10.1016/j.knee.2004.03.002.
38. Reilly KA, Beard DJ, Barker KL, et al. Efficacy of an accelerated recovery protocol for Oxford unicompartmental knee arthroplasty – a randomised controlled trial. *Knee* 2005; 12: 351-357, DOI: 10.1016/j.knee.2005.01.002.
39. Murray DW, Fitzpatrick R, Rogers K, et al. The use of the Oxford hip and knee scores. *J Bone Joint Surg Br* 2007; 89: 1010-1014, DOI: 10.1302/0301-620X.89B8.19424.
40. Weale AE, Halabi OA, Jones PW, White SH. Perceptions of outcomes after unicompartmental and total knee replacements. *Clin Orthop Relat Res* 2001; 382: 143-153, DOI: 10.1097/00003086-200101000-00021.
41. Haverkamp D, Breugem SJ, Sierevelt IN, et al. Translation and validation of the Dutch version of the Oxford 12-item knee questionnaire for knee arthroplasty. *Acta Orthop* 2005; 76: 347-352.
42. Dunbar MJ, Robertsson O, Ryd L, Lidgren L. Translation and validation of the Oxford-12 item knee score for use in Sweden. *Acta Orthop Scand* 2000; 71: 268-274, DOI: 10.1080/000164700317411861.
43. Martinelli N, Longo UG, Marinozzi A, et al. Cross-cultural adaptation and validation with reliability, validity, and responsiveness of the Italian version of the Oxford Hip Score in patients with hip osteoarthritis. *Qual life Res* 2011; 20: 923-929, DOI: 10.1007/s11136-010-9811-5.
44. Kavanagh BF, Fitzgerald RH Jr. Clinical and roentgenographic assessment of total hip arthroplasty. A new hip score. *Clin Orthop Relat Res* 1985; 193: 133-140.
45. Singh JA, Schleck C, Harmsen WS, Lewallen DG. Validation of the Mayo Hip Score: construct validity, reliability and responsiveness to change. *BMC Musculoskelet Disord* 2016; 17:39, DOI: 10.1186/s12891-016-0868-3.
46. Bremander AB, Petersson IF, Roos EM. Validation of the Rheumatoid and Arthritis Outcome Score (RAOS) for the lower extremity. *Health Qual Life Outcomes* 2003; 1: 55, DOI: 10.1186/1477-7525-1-55.
47. Lovelock TM, Broughton NS, Williams CM. The Popularity of Outcome Measures for Hip and Knee Arthroplasties. *J Arthroplasty* 2018; 33: 273-276, DOI: 10.1016/j.arth.2017.08.024.
48. Arden NK, Kiran A, Judge A, et al. What is a good patient reported outcome after total hip replacement? *Osteoarthritis Cartilage* 2011; 19: 155-162, DOI: 10.1016/j.joca.2010.10.004.
49. Berliner JL, Brodke DJ, Chan V, et al. John Charnley Award: Pre-operative Patient-reported Outcome Measures Predict Clinically Meaningful Improvement in Function After THA. *Clin Orthop Relat Res* 2016; 474: 321-329, DOI: 10.1007/s11999-015-4350-6.



50. Singh JA, Dowsey MM, Dohm M, et al. Achieving Consensus on Total Joint Replacement Trial Outcome Reporting Using the OMERACT Filter: Endorsement of the Final Core Domain Set for Total Hip and Total Knee Replacement Trials for Endstage Arthritis. *J Rheumatol* 2017; 44: 1723-1726, DOI: 10.3899/jrheum.161113.
51. Santić V, Legović D, Sestan B, et al. Measuring improvement following total hip and knee arthroplasty using the SF-36 Health Survey. *Coll Antropol* 2012; 36: 207-212.
52. Chang CH. Patient-reported outcomes measurement and management with innovative methodologies and technologies. *Qual Life Res* 2007; (16 Suppl 1): 157-166, DOI: 10.1007/s11136-007-9196-2.
53. Tariq MB, Vega JF, Westermann R, et al. Arthroplasty studies with greater than 1000 participants: analysis of follow-up methods. *Arthroplast Today* 2019; 5: 243-250, DOI: 10.1016/j.artd.2019.03.006.