

## Which is the Best Activity Rating Scale for Patients Undergoing Total Joint Arthroplasty?

Florian D. Naal MD, Franco M. Impellizzeri MS,  
Michael Leunig MD

Received: 13 February 2008 / Accepted: 9 June 2008 / Published online: 28 June 2008  
© The Association of Bone and Joint Surgeons 2008

**Abstract** We compared the metric properties of the University of California, Los Angeles (UCLA) activity scale, the Tegner score, and the Activity Rating Scale for assessment of activity levels in 105 patients undergoing THA (48 women; mean age, 63.4 years) and 100 patients undergoing TKA (61 women; mean age, 66.5 years). We assessed construct validity by correlating these scales with the International Physical Activity Questionnaire and different traditional patient self-reporting outcome measures. Test-retest reliability, feasibility, and floor and ceiling effects also were determined. The UCLA scale showed the strongest correlations with the other measures ( $r = -0.35$  to  $0.56$  for THA;  $r = -0.55$  to  $0.23$  for TKA) and was the only scale that discriminated between insufficiently and sufficiently active patients undergoing THA and TKA. The UCLA scale had the best reliability, provided the highest completion rate, and showed no floor effects. It seems to be the most

appropriate scale for assessment of physical activity levels in patients undergoing total joint arthroplasty.

**Level of Evidence:** Level III, diagnostic study. See the Guidelines for Authors for a complete description of levels of evidence.

### Introduction

Total joint arthroplasty (TJA) is one of the most successful interventions to reduce pain, improve joint function, and improve health-related quality of life in end-stage arthritic joints [15]. Pain relief and restoration of joint function still are the primary goals of surgery. However, patient expectations are increasing regarding physical activity and sports participation after surgery. In 1994, Wright et al. [42] reported participation in recreational activities was the third most important expectation of patients undergoing THA. In two other studies, Mont et al. [27, 28] reported some of their patients underwent TJA specifically to be able to continue playing tennis. Furthermore, participation in regular physical activity is important to improve cardiorespiratory fitness and reduce the morbidity and mortality associated with many chronic diseases [6].

However, despite widespread knowledge regarding the beneficial effects of physical activity on quality of life and general health enhancement [2], physical activity levels typically are not assessed in outcomes after TJA [4, 14]. Assessing physical activity levels is important given the negative consequences of activity in patients undergoing TJA: Kilgus et al. [22], and more recently Schmalzried et al. [34], described associations between physical activity levels and the risk of earlier implant failure. Flugsrud et al. [17] reported men with intermediate to intense physical activity during leisure time had a fourfold increased risk of

---

Each author certifies that he or she has no commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.

Each author certifies that his or her institution has approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.

---

F. D. Naal (✉), M. Leunig  
Department of Orthopaedic Surgery, Schulthess Clinic,  
Lengghalde 2, 8008 Zurich, Switzerland  
e-mail: florian.naal@gmail.com

F. M. Impellizzeri  
Neuromuscular Research Laboratory, Schulthess Clinic,  
Zurich, Switzerland

M. Leunig  
University of Berne, Berne, Switzerland

aseptic cup loosening compared with their counterparts with sedentary lifestyles. Unfortunately, as of now, there are no clear guidelines regarding what level of physical activity is sufficient to attain adequate cardiorespiratory and muscular fitness without compromising implant durability. Thus, measurement of patient activity is crucial in the evaluation of outcomes after TJA.

Determining patient activity is challenging and the assessment ideally should include the types of activities performed and the frequency, duration, and intensity of participation. There currently is no single approach addressing all these issues. Numerous physical activity questionnaires have been developed [7, 21, 26], but some lack validity [37] and others impose a heavy burden on patients and clinicians owing to their length and complexity [23, 37]. For assessment of physical activity, single rating scales asking the patient to rate her or his activity level on a numeric scale are shorter and more practical than the long questionnaires commonly used to assess habitual physical activity [23, 37]. For this reason, they easily can be used together with other traditional self-reported outcome measures in patients undergoing TJA. Activity rating scales, however, are not the ideal instruments to assess all issues relevant to physical activity. Nonetheless, as long as no gold standard exists, they probably should be used in the routine outcome assessment of TJA to address the above-mentioned issues of physical activity in this group of patients.

The purposes of our study were (1) to evaluate the validity of the UCLA activity scale [43], the Tegner score [39], and the Activity Rating Scale (ARS) [25] in patients undergoing TJA by correlating these scales with the International Physical Activity Questionnaire (IPAQ) [11] and different traditional outcome measures; (2) to compare the metric properties of these scales regarding reliability, feasibility, and floor and ceiling effects; and (3) to assess gender- and age-related effects on score values and metric properties.

## Materials and Methods

The study sample comprised 105 consecutive patients (48 women, 57 men) undergoing THA and 100 consecutive patients (61 women, 39 men) undergoing TKA between October and November 2007. Mean age of the THA group was  $63.4 \pm 11.0$  years (range, 33–88 years), with no differences ( $p = 0.16$ ) between women ( $65.1 \pm 10.5$  years) and men ( $62.0 \pm 11.2$  years). Mean body mass index (BMI) of the THA group was  $26.3 \pm 3.7$  kg/m<sup>2</sup> (range, 19.1–38.4 kg/m<sup>2</sup>). Mean age of the TKA group was  $66.5 \pm 9.1$  years (range, 46–88 years), with no differences ( $p = 0.22$ ) between women ( $67.6 \pm 8.3$  years) and men ( $65.2 \pm 10.3$  years). Mean BMI of the TKA group was

$28.1 \pm 3.9$  kg/m<sup>2</sup> (range, 19.5–41.2 kg/m<sup>2</sup>). All patients were sent a questionnaire set 1 week before surgery accompanied by a letter of explanation and instructions to complete the forms at home and bring them on the day of clinic admission. We changed the order of the activity scales in the questionnaire set to avoid repetitiveness bias. After receipt of the first questionnaire set, 43 patients (21 women, 22 men; mean age,  $63.4 \pm 10.3$  years; mean BMI,  $26.0 \pm 3.7$  kg/m<sup>2</sup>) from the THA group and 36 patients (18 women, 18 men; mean age,  $67.5 \pm 8.9$  years; mean BMI,  $27.8 \pm 4.0$  kg/m<sup>2</sup>) from the TKA group volunteered to complete a second questionnaire set to assess reliability. Evaluation of the activity scales was part of a large validation study approved by the local ethical committee. All participating patients provided written informed consent.

We used the UCLA scale [43], Tegner score [39], and ARS [25]. The UCLA scale is a simple scale ranging from 1 to 10. The patient indicates her or his most appropriate activity level, with 1 defined as “no physical activity, dependent on others” and 10 defined as “regular participation in impact sports.” The Tegner score is similar to the UCLA scale, a simple scale ranging from 0 to 10, and the patient has to indicate the most appropriate activity level, with 0 defined as “no physical activity, disabled” and 10 defined as “participation in competitive soccer—national and international elite.” The ARS consists of four questions asking about the frequency the patient performs activities such as “running, cutting, decelerating, and pivoting.” Each question can be scored from 0 (less than one time per month) to 4 (four times per week or more often), so the total ARS can range from 0 to 16 points.

No gold standard questionnaire exists to assess physical activity. We chose the short “last 7 days” version of the International Physical Activity Questionnaire (IPAQ; [www.ipaq.ki.se](http://www.ipaq.ki.se)) because it is considered valid [11] and its use in scientific studies is widespread [1, 31, 32, 35, 40]. This IPAQ version consists of seven questions assessing the frequency and duration of participation in vigorous, moderate, and walking activity, and the time spent sitting during the last week. The total score can be expressed in metabolic equivalents (METs). One MET is approximately the metabolic rate (oxygen consumption) of an individual sitting quietly for 1 minute (3.5 mL/kg/minute). To estimate total METs, we multiplied total minutes spent in each category per week by a factor of 8 for vigorous, 4 for moderate, and 3.3 for walking activity, as proposed by the IPAQ guidelines. Data truncation and removal of outliers also were performed, as proposed by the IPAQ guidelines; 19% of the questionnaires were truncated, and one outlier (undergoing THA) had to be removed.

To assess construct validity, we used two approaches: convergent evidence [3] and known-group technique [3]. Convergent evidence was determined by correlating the

activity rating scales with total METs calculated from the IPAQ. Because our goal was to evaluate physical activity rating scales in patients undergoing THA or TKA, we presumed convergent evidence between the scores of the activity rating scales and the scores derived from generic and disease-specific questionnaires traditionally used to assess surgical outcomes in patients undergoing TJA: WOMAC [5], Oxford Hip Score [12] (OHS), SF-12 [41], and Harris hip score [18] (HHS) in patients undergoing THA; and WOMAC, Oxford Knee Score [13] (OKS), SF-12, and Knee Society score [20] (KSS) with knee and function subscores in patients undergoing TKA. These relations were analyzed using the Spearman rank correlation coefficient. We anticipated there would be weak to moderate correlations ( $r = 0.3\text{--}0.5$ ) between the activity scales and the generic and disease-specific measures.

The second procedure we used for gathering construct-related evidence was the known-group difference technique in which two groups assumed to differ on the construct being measured are compared. Therefore, according to the guidelines set out by the IPAQ executive committee, patients were classified as insufficiently active, moderately active, or vigorously active. We calculated the mean score values of the activity scales for patients classified as insufficiently active and sufficiently active (moderate and vigorous activity levels according to the IPAQ classification), and we then assessed if these values differed between the groups using the Mann-Whitney U test.

The reliability of the three activity rating scales (UCLA, Tegner, ARS) was examined using the weighted Cohen's kappa  $k_w$  [10]. The coefficient  $k$  indicates whether observed agreement is greater than or equal to chance agreement. Before calculation, each cell of the contingency table was assigned weights. Cells with exact agreement were assigned the maximum of 1. The quadratic weights from these maximum values ranged down to 0 for cells with a maximal possible disagreement [16]. The strength of agreement for the  $k_w$  coefficient was interpreted according to Fleiss et al. [16]:  $k_w$  of 0.75 or more indicates excellent agreement for most purposes and  $k_w$  of 0.40 or less indicates poor agreement. The test-retest reliability of IPAQ was examined using the intraclass correlation coefficient [ICC(2,1): two-way random-effects model with single measure (absolute agreement)] [38].

We calculated the rate of fully completed activity scales. Furthermore, each scale had to be rated by patients for its difficulty on a visual analog scale from 1 (very easy) to 10 (extremely difficult), and patients stated the time they required for completion. We compared all three rating scales regarding difficulty and time consumption using the Kruskal-Wallis test with post hoc Mann-Whitney U tests. Both parameters also were correlated with patients' age using Spearman rank correlation coefficients.

For all three scales, the distribution of floor and ceiling effects was calculated. Floor effects occur when patients rate their activity level as the lowest possible on the specific scale (1 for the UCLA scale, 0 for the ARS and Tegner scale). Consequently, worsening in activity levels cannot be assessed. Vice versa, ceiling effects occur when patients rate their activity level as the highest possible on the specific scale (10 for the UCLA and Tegner, 16 for the ARS). Ceiling effects therefore prohibit observing any improvement in activity levels.

Before all statistical tests, we tested the data for normal distribution using the Shapiro-Wilk W test. Unless otherwise stated, descriptive results are shown as mean  $\pm$  standard deviation. For the reliability values, we indicated 95% confidence intervals (CIs). All statistical tests were performed using the software package SPSS (Version 13; SPSS Inc, Chicago, IL).

## Results

The three scales showed weak to moderate correlations with the other scores used in the expected directions; we found the strongest correlations between activity scales and IPAQ, WOMAC, OHS, OKS, HHS, KSS, and SF-12 values for the UCLA scale in patients undergoing THA and those undergoing TKA (Table 1). We observed the weakest correlations for the ARS (Table 1). The UCLA scale discriminated between insufficiently and sufficiently active patients as classified by the IPAQ in patients undergoing THA and those undergoing TKA (Table 2). The ARS could discriminate only between insufficiently and sufficiently active patients undergoing THA but not in patients undergoing TKA (Table 2). Similarly, the Tegner could discriminate only between insufficiently and sufficiently active patients in the THA group but not in the TKA group (Table 2).

The UCLA scale had excellent reliability with higher  $k_w$  values than the ARS and Tegner score in patients undergoing THA (Table 3). The reliability of all three scales was excellent in patients undergoing TKA (Table 3). The ICCs for the IPAQ were 0.76 [95% CI, 0.57–0.87] in patients undergoing THA and 0.87 [95% CI, 0.74–0.94] in patients undergoing TKA.

The UCLA scale provided the highest completion rate in patients undergoing THA and those undergoing TKA (Table 4). We observed no major differences between all scales concerning difficulty and time required for completion (Table 4). No floor effects were observed for the UCLA scale in either the THA or the TKA group. Ceiling effects occurred in 4% of the patients undergoing THA and in 1% of the patients undergoing TKA. We observed large floor effects for the ARS: 56% for patients undergoing

**Table 1.** Validity of the activity rating scales

Scale	Patients undergoing THA (n = 105)						Patients undergoing TKA (n = 100)					
	UCLA		ARS		Tegner		UCLA		ARS		Tegner	
	r value	p level	r value	p level	r value	p level	r value	p level	r value	p level	r value	p level
IPAQ total	0.52	< 0.0001	0.29	0.0045	0.39	< 0.0001	0.23	0.0382	0.30	0.0210	0.20	0.0744
IPAQ sitting	-0.04	0.7032	0.03	0.7663	0.08	0.4297	-0.25	0.0124	-0.28	0.0135	-0.11	0.3047
WOMAC total	-0.49	< 0.0001	-0.26	0.0045	-0.39	< 0.0001	-0.50	< 0.0001	-0.26	0.0148	-0.39	0.0001
WOMAC pain	-0.45	< 0.0001	-0.24	0.0087	-0.32	0.0004	-0.47	< 0.0001	-0.28	0.0086	-0.39	0.0001
WOMAC stiffness	-0.35	< 0.0001	-0.30	0.0008	-0.40	< 0.0001	-0.24	0.0113	-0.14	0.1869	-0.24	0.0172
WOMAC function	-0.49	< 0.0001	-0.25	0.0069	-0.39	< 0.0001	-0.50	< 0.0001	-0.26	0.0147	-0.37	0.0002
SF-12 PCS	0.53	< 0.0001	0.28	0.0034	0.37	0.0001	0.46	< 0.0001	0.25	0.0123	0.31	0.0016
SF-12 MCS	0.38	< 0.0001	0.08	0.3842	0.24	0.0147	0.47	< 0.0001	0.10	0.3429	0.05	0.6227
OHS	-0.48	< 0.0001	-0.25	0.0048	-0.42	< 0.0001	NA		NA		NA	
OKS	NA		NA		NA		-0.55	< 0.0001	-0.29	0.0174	-0.29	0.0039
HHS	0.56	< 0.0001	0.29	0.0150	0.39	0.0027	NA		NA		NA	
KSS total	NA		NA		NA		0.51	< 0.0001	0.43	0.0013	0.14	0.3214
KSS knee score	NA		NA		NA		0.33	0.0069	0.30	0.0320	-0.06	0.6449
KSS function score	NA		NA		NA		0.51	< 0.0001	0.41	0.0026	0.38	0.0047

The UCLA scale correlated the strongest with other measures used in patients undergoing THA and those undergoing TKA; correlation coefficients were lower for Tegner and ARS; UCLA = University of California, Los Angeles activity scale; ARS = Activity Rating Scale; IPAQ = International Physical Activity Questionnaire; PCS = Physical Component Summary; MCS = Mental Component Summary; OHS = Oxford Hip Score; OKS = Oxford Knee Score; HHS = Harris hip score; KSS = Knee Society score; NA = not available.

**Table 2.** Activity scale values of sufficiently and insufficiently active patients

Scale	Patients undergoing THA (n = 105)			Patients undergoing TKA (n = 100)		
	Sufficiently active (42%)	Insufficiently active (58%)	p level	Sufficiently active (67%)	Insufficiently active (33%)	p level
UCLA	5.8 ± 1.9 (5.3–6.4)	4.7 ± 2.1 (4.2–5.1)	0.0005	5.3 ± 2.0 (4.9–5.8)	4.4 ± 2.0 (3.8–5.1)	0.0139
ARS	4.3 ± 4.7 (2.9–5.7)	2.4 ± 4.0 (1.5–3.2)	0.0127	3.0 ± 4.4 (1.8–4.1)	2.3 ± 3.5 (1.0–3.7)	0.7967
Tegner	3.1 ± 1.7 (2.6–3.6)	2.4 ± 1.2 (2.1–2.6)	0.0093	2.6 ± 1.5 (2.3–3.0)	2.5 ± 2.0 (1.8–3.3)	0.2876

Values are mean ± standard deviation, with 95% confidence intervals in parentheses; UCLA = University of California, Los Angeles activity scale; ARS = Activity Rating Scale.

**Table 3.** Reliability of the activity rating scales

Scale	Weighted kappa (k <sub>w</sub> )	
	Patients undergoing THA (n = 43)	Patients undergoing TKA (n = 36)
UCLA	0.80 (0.70–0.90)	0.86 (0.77–0.95)
ARS	0.65 (0.47–0.84)	0.88 (0.77–0.99)
Tegner	0.54 (0.24–0.83)	0.84 (0.68–0.99)

95% confidence intervals in parentheses; UCLA = University of California, Los Angeles activity scale; ARS = Activity Rating Scale.

THA and 55% for patients undergoing TKA. Ceiling effects occurred in 1% and 2%, respectively. Floor effects occurred in 6% for the patients undergoing THA and 3% for the patients undergoing TKA for the Tegner scale. The Tegner scale showed no ceiling effects.

According to all three scales, men were more active (p = 0.000004–0.006) than women (Table 5). In patients

undergoing THA, we observed no correlation (p = 0.08–0.95) between age and the activity scales or IPAQ values. Similarly, in patients undergoing TKA, we observed no correlation (p = 0.25–0.97) between age and the activity scales or IPAQ values. In patients undergoing THA, the time required to complete the UCLA scale and ARS increased (r = 0.21, p = 0.046, and r = 0.29, p = 0.01, respectively) with age. We observed no association (p = 0.10) between time and age for completion of the Tegner scale. Difficulty in completing the scales was not correlated (p = 0.11–0.57) with age in patients undergoing THA. For patients undergoing TKA, the time required to complete the ARS was correlated (r = 0.28, p = 0.02) with age. We found no association between time and age for the UCLA scale (p = 0.09) and Tegner scale (p = 0.10). The difficulty for patients undergoing TKA to complete the ARS and Tegner scale increased (r = 0.24, p = 0.05, and r = 0.25, p = 0.03, respectively) with age.

**Table 4.** Feasibility of the activity rating scales

Scale	Patients undergoing THA (n = 105)			Patients undergoing TKA (n = 100)		
	Completed	Time (minutes)	Difficulty	Completed	Time (minutes)	Difficulty
UCLA	95%	3.8 ± 0.6	2.4 ± 0.3	93%	3.9 ± 0.5*	2.5 ± 0.3
ARS	86%	3.2 ± 0.4	2.8 ± 0.4	71%	2.4 ± 0.3	2.6 ± 0.6
Tegner	82%	4.2 ± 0.7	2.8 ± 0.4	76%	3.3 ± 0.6	3.0 ± 0.5
IPAQ	92%	NA	NA	84%	NA	NA

Values are mean ± standard deviation; \* $p = 0.032$ ; highest completion rates were found for the UCLA; patients undergoing TKA required more ( $p = 0.032$ ) time to complete the UCLA than the ARS, but after removal of scales with floor and ceiling effects this difference lost statistical significance ( $p = 0.18$ ); there were no differences between all scales concerning the difficulty to complete as assessed by a visual analog scale (from 1 to 10); UCLA = University of California, Los Angeles activity scale; ARS = Activity Rating Scale; IPAQ = International Physical Activity Questionnaire; NA = not available.

We observed no association ( $p = 0.68$ ) between difficulty and age for completion of the UCLA scale.

## Discussion

Assessing physical activity levels in patients undergoing TJA is important considering the associations between physical activity levels and the risk of earlier implant failure [17, 22, 34]. However, physical activity levels are not routinely determined in the outcome assessment after TJA [4, 14]. Using activity rating scales is one possibility to assess activity levels in patients undergoing TJA, and as long as no gold standard exists, they probably should be used in the routine outcome assessment after THA or TKA. Our purposes were to evaluate the validity of the UCLA activity scale [43], the Tegner score [39], and the ARS [25] in patients undergoing TJA; to compare the metric properties of these scales regarding reliability, feasibility, and floor and ceiling effects; and to assess gender- and age-related effects on score values and metric properties.

When interpreting the results, one must consider we investigated only three different activity rating scales, and we used the IPAQ as a reference for physical activity and no objective measure, such as a pedometer or accelerometer.

Our data showed, in patients undergoing TJA, the UCLA activity scale correlated better with the other measures, and provided better reliability and completion rate than the Tegner scale and the ARS. The UCLA scale showed consistent construct validity and it was the only scale discriminating between insufficiently and sufficiently active patients, although the 95% CIs overlapped in the TKA group. Therefore, of these three scales, the UCLA seems most appropriate for assessing activity levels in patients undergoing TJA. Our data are consistent with those of two previous studies reporting correlation coefficients between the UCLA scale, HHS, and the SF-12 [4, 43]. The correlation coefficients we observed between the UCLA scale and the IPAQ compared well with those

reported between the IPAQ and accelerometer data [11], indicating construct validity of the UCLA scale. The UCLA scale and Tegner score seemed to perform better in patients undergoing THA than in patients undergoing TKA, considering higher correlation coefficients with the IPAQ. However, we attribute this observation not to the UCLA scale or Tegner score but to the IPAQ: according to all three activity scales used in this study, patients undergoing THA had higher activity levels than patients undergoing TKA, but in contrast, total METs derived from the IPAQ were higher in patients undergoing TKA, and more patients undergoing TKA than THA were classified as sufficiently active. This contradictory pattern can explain the lower correlation coefficients found in the TKA group, and might indicate overreporting of habitual physical activity in the TKA group using the IPAQ. The possible problem of overreporting with the IPAQ was highlighted by others [1, 33, 40]. Whether overreporting physical activity is responsible for the observed differences and why this seemed to occur only in patients undergoing TKA should be investigated in future studies using the IPAQ or other habitual physical activity questionnaires together with objective measures, such as pedometers or accelerometers, in patients undergoing THA and TKA.

We observed a high rate of floor effects using the ARS. Floor effects were observed in greater than 50%. Moreover, the ARS had the lowest completion rate and the weakest correlations with the other measures used in this study. The ARS originally was developed to assess activity levels in patients with knee injuries [25]. Such patients are younger and might have different demands on their knee function. This may explain floor effects and the low completion rate; however, our results indicate this scale should not be used in patients undergoing TJA. Similar to the ARS, the Tegner scale was developed to rate activity levels in patients with knee injuries [39]. The better metric properties achieved by the Tegner scale in comparison to the ARS might be explained by its design as a simple scale from 0 to 10, which is similar to the UCLA scale design. Despite the

**Table 5.** Descriptive results for the activity rating scales

Scale	Patients undergoing THA			Patients undergoing TKA			
	All (n = 105)	Women (n = 48)	Men (n = 57)	All (n = 100)	Women (n = 61)	Men (n = 39)	p level
UCLA	5.1 ± 2.1 (2–10)	4.4 ± 1.8 (2–10)	5.7 ± 2.2 (2–10)	4.9 ± 2.0 (2–10)	4.5 ± 1.7 (2–9)	5.6 ± 2.2 (2–10)	0.0058
ARS	3.0 ± 4.3 (0–16)	1.6 ± 3.7 (0–14)	4.0 ± 4.5 (0–16)	2.5 ± 3.9 (0–16)	1.6 ± 3.0 (0–12)	3.7 ± 4.6 (0–16)	0.0045
Tegner	2.7 ± 1.5 (0–7)	2.4 ± 1.5 (0–6)	2.8 ± 1.5 (0–7)	2.6 ± 1.6 (0–9)	2.0 ± 1.5 (0–9)	3.2 ± 1.6 (1–7)	< 0.0001
IPAQ (MET/minutes/week <sup>-1</sup> )	3690 ± 5168 (0–38556)	3501 ± 4184 (0–20112)	3844 ± 5879 (0–38556)	4623 ± 6508 (0–42660)	3826 ± 3930 (0–16212)	5592 ± 8619 (0–42660)	0.9081

Values are mean ± standard deviation, with range in parentheses; according to all three activity rating scales, men were more active than women; UCLA = University of California, Los Angeles activity scale; ARS = Activity Rating Scale; IPAQ = International Physical Activity Questionnaire; MET = metabolic equivalents.

similar design, the Tegner scale had lower correlation coefficients, lower reliability values, and lower completion rates than the UCLA scale. This might be related to a too sports-oriented content for the group of patients undergoing TJA; Levels 8 to 10 are related to competitive sports participation. With few exceptions, the majority of patients do not undergo joint arthroplasty to achieve competitive sports levels [24, 42]. The rate of floor effects we found for the Tegner scale was less than 10%, which might be considered an acceptable proportion, bearing in mind these patients were undergoing TJA.

All three activity scales showed men had higher activity levels than women. These gender-related differences were in line with previous reports. Ainsworth et al. [1] and Tehard et al. [40] reported men were more active than women according to the IPAQ. Using a pedometer to assess physical activity, Sequeira et al. [36] reported Swiss women were less active than men. We found no age-dependent decrease in physical activity levels, contrary to the findings of Zahiri et al. [43] using the UCLA scale. This observation seems to indicate older adults undergoing TJA are still healthy and active, which underlines the importance to assess physical activity levels in the outcome assessment of joint arthroplasty. Otherwise, this observation also might be related to geographic differences between US and Swiss patients because Zahiri et al. [43] also investigated patients undergoing THA.

The UCLA scale has been criterion-validated examining the correlations between UCLA scores and objective measures of physical activity obtained with pedometers [43]. Criterion validity has, to our knowledge, not yet been proven for the ARS or Tegner scale. Nevertheless, one major weakness of the UCLA scale is this scale, similar to the Tegner scale and ARS, does not assess frequency, duration, and intensity of activities. Considering the benefits of health-enhancing physical activity levels and the risks of repetitive load cycles applied on a prosthesis during physical activities, assessment of frequency, duration, and intensity of activities is of utmost importance. There is increasing evidence prosthetic wear is not a function of time but of use [34]. Moreover, patient expectations concerning physical activity levels after TJA are increasing. In 1994, Wright et al. [42] reported the wish to return to recreational activities such as golfing or gardening was the third most important concern for patients undergoing TJA. This concern was cited by 85% of the patients. Similarly, Mancuso et al. [24] reported return to sports was a major expectation of surgery in their patients undergoing knee surgery. Mont et al. [27, 28] investigated the ability to play tennis after THA and TKA and reported some of their patients underwent surgery specifically to be able to continue playing tennis. The majority of patients will participate in recreation and sports after surgery [8, 9, 19].

The proportion of physically and sports-active patients after TJA varies among reports but may exceed 90% [29, 30]. Therefore, assessment of activity levels in patients undergoing hip and knee arthroplasty is important to respect patients' concerns and expectations and to cover possible risks derived from increased physical activity and sports concerning implant failure and wear production.

The UCLA scale is a reliable, feasible, and valid instrument for assessment of activity levels in patients undergoing TJA. The Tegner scale has inferior metric properties compared with the UCLA scale, and the ARS should not be used in this group of patients. Nevertheless, simple measures are needed to assess the frequency, duration, and intensity of performed activities. Future studies should investigate whether simple activity measures such as the UCLA scale allow for assessment of long-term relationships between activities and implant failure.

**Acknowledgments** We thank Markus Loibl and Marc Sieverding for assistance with this study.

## References

- Ainsworth BE, Macera CA, Jones DA, Reis JP, Addy CL, Bowles HR, Kohl HW 3rd. Comparison of the 2001 BRFSS and the IPAQ Physical Activity Questionnaires. *Med Sci Sports Exerc.* 2006;38:1584–1592.
- American College of Sports Medicine Physicians Statement. The recommended quantity and quality of exercise for developing and maintaining cardiorespiratory and muscular fitness in healthy adults. *Med Sci Sports Exerc.* 1990;22:265–274.
- Ary D, Cheser Jacobs L, Razavieh A, Sorensen C. *Introduction to Research in Education.* 6th ed. Belmont, CA: Wadsworth; 2006.
- Beaulé PE, Dorey FJ, Hoke R, Leduff M, Amstutz HC. The value of patient activity level in the outcome of total hip arthroplasty. *J Arthroplasty.* 2006;21:547–552.
- Bellamy N, Buchanan WW, 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 osteoarthritis of the hip or knee. *J Rheumatol.* 1988;15:1833–1840.
- Blair SN, Kohl HW 3rd, Barlow CE, Pfaffenbarger RS Jr, Gibbons LW, Macera CA. Changes in physical fitness and all-cause mortality: a prospective study of healthy and unhealthy men. *JAMA.* 1995;273:1093–1098.
- Bonnefoy M, Normand S, Pachiardi C, Lacour JR, Laville M, Kostka T. Simultaneous validation of ten physical activity questionnaires in older men: a doubly labeled water study. *J Am Geriatr Soc.* 2001;49:28–35.
- Bradbury N, Borton D, Spoo G, Cross MJ. Participation in sports after total knee replacement. *Am J Sports Med.* 1998;26:530–535.
- Clifford PE, Mallon WJ. Sports after total joint replacement. *Clin Sports Med.* 2005;24:175–186.
- Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull.* 1968;70:213–220.
- Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, Pratt M, Ekelund U, Yngve A, Sallis JF, Oja P. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc.* 2003;35:1381–1395.
- 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.
- Dawson J, Fitzpatrick R, Murray D, Carr A. Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br.* 1998;80:63–69.
- Dorey FJ, Amstutz HC. The need to account for patient activity when evaluating the results of total hip arthroplasty with survivorship analysis. *J Bone Joint Surg Am.* 2002;84:709–710.
- Ethgen O, Bruyère O, Richey F, Dardennes C, Reginster JY. Health-related quality of life in total hip and total knee arthroplasty: a qualitative and systematic review of the literature. *J Bone Joint Surg Am.* 2004;86:963–974.
- Fleiss JL, Levin B, Paik MC. *Statistical Methods for Rates, Proportions.* 3rd ed. Hoboken, NJ: John Wiley and Sons, Inc; 2003.
- Flugsrud GB, Nordsletten L, Espehaug B, Havelin LI, Meyer HE. The effect of middle-age body weight and physical activity on the risk of early revision hip arthroplasty: a cohort study of 1,535 individuals. *Acta Orthop.* 2007;78:99–107.
- 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.
- Healy WL, Iorio R, Lemos MJ. Athletic activity after joint replacement. *Am J Sports Med.* 2001;29:377–388.
- Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clin Orthop Relat Res.* 1989;248:13–14.
- Jacobs DR Jr, Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Med Sci Sports Exerc.* 1993;25:81–91.
- Kilgus DJ, Dorey FJ, Finerman GA, Amstutz HC. Patient activity, sports participation, and impact loading on the durability of cemented total hip replacements. *Clin Orthop Relat Res.* 1991;269:25–31.
- Kriska AM, Caspersen CJ. Introduction to a collection of physical activity questionnaires. *Med Sci Sports Exerc.* 1997;29(suppl):5–9.
- Mancuso CA, Sculco TP, Wickiewicz TL, Jones EC, Robbins L, Warren RF, Williams-Russo P. Patients' expectations of knee surgery. *J Bone Joint Surg Am.* 2001;83:1005–1012.
- Marx RG, Stump TJ, Jones EC, Wickiewicz TL, Warren RF. Development and evaluation of an activity rating scale for disorders of the knee. *Am J Sports Med.* 2001;29:213–218.
- Miller DJ, Freedson PS, Kline GM. Comparison of activity levels using the Caltrac accelerometer and five questionnaires. *Med Sci Sports Exerc.* 1994;26:376–382.
- Mont MA, LaPorte DM, Mullick T, Silberstein CE, Hungerford DS. Tennis after total hip arthroplasty. *Am J Sports Med.* 1999;27:60–64.
- Mont MA, Rajadhyaksha AD, Marxen JL, Silberstein CE, Hungerford DS. Tennis after total knee arthroplasty. *Am J Sports Med.* 2002;30:163–166.
- Naal FD, Fischer M, Preuss A, Goldhahn J, von Knoch F, Preiss S, Munzinger U, Drobny T. Return to sports and recreational activities after unicompartmental knee arthroplasty. *Am J Sports Med.* 2007;35:1688–1695.
- Naal FD, Maffiuletti NA, Munzinger U, Hersche O. Sports after hip resurfacing arthroplasty. *Am J Sports Med.* 2007;35:705–711.
- Rosemann T, Kuehlein T, Laux G, Szecsenyi J. Osteoarthritis of the knee and hip: a comparison of factors associated with physical activity. *Clin Rheumatol.* 2007;26:1811–1817.
- Rütten A, Ziemainz H, Schena F, Stahl T, Stiggelbout M, Auweele YV, Vuillemin A, Welshman J. Using different physical

- activity measurements in eight European countries: results of the European Physical Activity Surveillance System (EUPASS) time series survey. *Public Health Nutr.* 2003;6:371–376.
33. Rzewnicki R, Vanden Auweele Y, De Bourdeaudhuij I. Addressing overreporting on the International Physical Activity Questionnaire (IPAQ) telephone survey with a population sample. *Public Health Nutr.* 2003;6:299–305.
  34. Schmalzried TP, Shepherd EF, Dorey FJ, Jackson WO, dela Rosa M, Fa'vae F, McKellop HA, McClung CD, Martell J, Moreland JR, Amstutz HC. The John Charnley Award. Wear is a function of use, not time. *Clin Orthop Relat Res.* 2000;381:36–46.
  35. Schmidt MD, Cleland VJ, Thomson RJ, Dwyer T, Venn AJ. A comparison of subjective and objective physical activity and fitness in identifying associations with cardiometabolic risk factors. *Ann Epidemiol.* 2008;18:378–386.
  36. Sequeira MM, Rickenbach M, Wietlisbach V, Tullen B, Schutz Y. Physical activity assessment using a pedometer and its comparison with a questionnaire in a large population survey. *Am J Epidemiol.* 1995;142:989–999.
  37. Shepard RJ. Limits to the measurement of habitual physical activity by questionnaires. *Br J Sports Med.* 2003;37:197–206; discussion 206.
  38. Shrout PE, Fleiss J. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;86:420–428.
  39. Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. *Clin Orthop Relat Res.* 1985;198:43–49.
  40. Tehard B, Saris WH, Astrup A, Martinez JA, Taylor MA, Barbe P, Richterova B, Guy-Grand B, Sørensen TI, Oppert JM. Comparison of two physical activity questionnaires in obese subjects: the NUGENOB study. *Med Sci Sports Exerc.* 2005;37:1535–1541.
  41. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care.* 1996;34:220–233.
  42. Wright JG, Rudicel S, Feinstein AR. Ask patients what they want: evaluation of individual complaints before total hip replacement. *J Bone Joint Surg Br.* 1994;76:229–234.
  43. Zahiri CA, Schmalzried TP, Szuszczewicz ES, Amstutz HC. Assessing activity in joint replacement patients. *J Arthroplasty.* 1998;13:890–895.