

# Cross-cultural adaptation, reliability and validity of the Japanese version of the Hospital for Special Surgery Hip Replacement Expectations Survey

薬師寺, 佳菜子

<https://doi.org/10.15017/4060018>

---

出版情報 : Kyushu University, 2019, 博士 (看護学), 課程博士  
バージョン :

権利関係 : © 2019 The Japanese Orthopaedic Association. Published by Elsevier B.V. All rights reserved.





Contents lists available at ScienceDirect

Journal of Orthopaedic Science

journal homepage: <http://www.elsevier.com/locate/jos>

## Original Article

# Cross-cultural adaptation, reliability and validity of the Japanese version of the Hospital for Special Surgery Hip Replacement Expectations Survey

Kanako Yakushiji<sup>a</sup>, Kimie Fujita<sup>a,\*</sup>, Satomi Tanaka<sup>a</sup>, Rumi Tanaka<sup>a</sup>, Jun-ichi Fukushi<sup>b</sup>, Goro Motomura<sup>b</sup>, Satoshi Hamai<sup>b</sup>, Satoshi Ikemura<sup>b</sup>, Masanori Fujii<sup>b</sup>, Yasuharu Nakashima<sup>b</sup>

<sup>a</sup> Department of Health Sciences, Faculty of Medical Sciences, Kyushu University, 3-1-1, Maidashi, Fukuoka City, Fukuoka, 812-8582, Japan

<sup>b</sup> Department of Orthopedic Surgery, Graduate School of Medical Sciences, Kyushu University, 3-1-1 Maidashi, Higashi-ku, Fukuoka, 812-8582, Japan

## ARTICLE INFO

## Article history:

Received 26 April 2019

Received in revised form

7 July 2019

Accepted 17 July 2019

Available online xxx

## ABSTRACT

**Background:** The Hospital for Special Surgery Hip Replacement Expectations Survey (HSS-THRES) is used in many countries to determine patient expectations before surgery. This study aimed to assess the reliability and validity of the Japanese version of HSS-THRES.

**Methods:** A total of 134 patients scheduled for total hip arthroplasty (THA) underwent a self-administered preoperative survey questionnaire. Patient's expectation and quality of life (QOL) were measured using the Japanese version of HSS-THRES, overall expectations for THA, Oxford hip score (OHS), and EuroQol-5D (EQ-5D). Some patients completed the Japanese version of HSS-THRES and the overall expectations for THA after a ten-day interval. Cross-cultural adaptation was validated by an expert committee comprising health professionals, a methodologist, language experts, and orthopedic specialists. The internal consistency was evaluated by the Cronbach  $\alpha$  coefficient. The test-retest reliability was examined using the intraclass coefficient correlation (ICC) and the Bland and Altman analysis. To test the construct validity, nine priori hypotheses were tested by correlation analysis between the Japanese version of HSS-THRES and two QOL scales, and by examining the association with demographic variables.

**Results:** A total of 116 patients completed four scales. Patients were predominantly female (75.9%), with an average age of  $62.2 \pm 11.7$ . In the cross-cultural adaptation, all patients responded to the questionnaire without problems. The Japanese version of HSS-THRES showed good internal consistency (Cronbach  $\alpha$ : 0.9). ICC was 0.94 and Bland–Altman analysis indicated no bias. The correlation between Japanese HSS-THRES and overall expectations for THA was high ( $r = 0.67$ ). Similarly, the correlation with the OHS was higher than that with EQ-5D. A total of 77.8% of the hypotheses were confirmed.

**Conclusions:** The Japanese version of HSS-THRES showed good cultural acceptability, high reliability, and validity to evaluate preoperative expectations for THA patients.

© 2019 The Japanese Orthopaedic Association. Published by Elsevier B.V. All rights reserved.

## 1. Introduction

Medical outcomes have focused not only on physician's outcome and evaluation, but also on patient-reported outcomes, such as health-related quality of life (HRQOL) and satisfaction [1,2]. When

there are several treatment options, it is necessary to grasp the patient's expectation to promote patient decision making [3]. The relation of expectation to total hip arthroplasty (THA) has been reported; high preoperative expectations and fulfilled expectations postoperatively were related to greater improvement in physical functions, better HRQOL, and higher satisfaction with the surgical results [4,5]. Thus, patient's expectation is an important consideration in the functional outcome of surgeries [5,6].

Unrealistic high expectations may weaken the doctor–patient trust and cause patients to be discouraged postoperatively [7],

\* Corresponding author. Department of Health Sciences, Faculty of Medical Sciences, Kyushu University, 3-1-1, Maidashi, Fukuoka City, Fukuoka, 812-8582, Japan. Fax: +92 642 6738.

E-mail address: [fujitak@hs.med.kyushu-u.ac.jp](mailto:fujitak@hs.med.kyushu-u.ac.jp) (K. Fujita).

<https://doi.org/10.1016/j.jos.2019.07.006>

0949-2658/© 2019 The Japanese Orthopaedic Association. Published by Elsevier B.V. All rights reserved.

whereas unrealistic low expectations can result in less motivation for treatment and rehabilitation [8]. To avoid such unrealistic expectations, it is important that the health providers grasp the patient's expectation preoperatively [4]. In addition, preoperative education classes have resulted in more equal expectations of both patient and surgeon [9]. A scale evaluating expectations can be used to evaluate interventions and to guide discussions between patients and surgeon.

A systematic review of THA patient's expectations reported that evaluation methods were different [10,11]. There is a bias in expectation survey items; the items related to pain and function in the lower limbs have been measured in most studies, but those related to activities of daily living, leisure activities, and walking assistive device were not [11]. However, recently the Hospital for Special Surgery Hip Replacement Expectations Survey (HSS-THRES) was developed to determine patient expectations before THA [4,12]. Its reliability and validity has been assessed in Germany [13], France [14], Holland [15], and China [16]. As developed by interviewing 180 patients about their expectations and reviewing of these patient-derived items by a panel of orthopedic surgeons, the items in the scale varied. The items are related to pain, walking, essential activities (work/sleeping/sexual activity, and so on), and psychological well-being. Therefore, developing the Japanese version of HSS-THRES would be useful for evaluating the preoperative expectations of Japanese THA patients and the intervention of medical providers.

The purpose of this study was to assess the reliability and validity of the Japanese version of HSS-THRES.

## 2. Materials and methods

### 2.1. The scales

#### 2.1.1. Hospital for Special Surgery Hip Replacement Expectations Survey (HSS-THRES)

HSS-THRES is a scale developed by Mancuso et al. to determine patients' expectations before THA [4,12]. It consists of 18 items, such as pain in the lower limbs, walking, and essential activities of daily living. Patients were asked how much relief or improvement they expected for each item as a result of their hip replacement. The patients rated their expectations using a five-point Likert scale (4 = complete improvement or back to normal and 0 = this expectation does not apply to me/I do not have this expectation). The total score ranged from 0 to 72 points, which was converted to 100 full points, with high scores representing high expectations. The original English version of the survey showed good test-retest reliability and content validity [4,9]. Cronbach's alpha as a measure of internal consistency was 0.77 [9].

#### 2.1.2. Overall expectation for THA

Patients were tested for the overall expectations for THA using the following question, "What is your overall expectation of the surgery?", the response for which was scored using a visual analog scale (0 = I do not have any expectations at all and 10 = I have great expectations) [17].

#### 2.1.3. Disease-specific HRQOL scale

The Oxford Hip Score (OHS) comprises 12 items addressing physical pain and function, with the total score ranging from 12 to 60. Higher scores were associated with poorer HRQOL. Of the 12 items, the score was calculated by dividing it into items related to pain (6 items) and those related to physical function (6 items) [18]. The Japanese version of the OHS was validated in 2009 [19].

#### 2.1.4. General HRQOL scale

The EuroQoL-5 Dimension (EQ-5D), a generic instrument used to assess QOL, comprises five items related to mobility, self-care, usual activity, pain/discomfort, and anxiety/depression. There are five possible response levels for each item. EQ-5D utility is a single weighted utility score calculated using the five items. Perfect health has a utility score of one, death has a score of zero, and states worse than death was rated <0. Higher EQ-5D item scores indicated worse outcomes, whereas higher EQ-5D utility scores indicated better health outcomes [20].

### 2.2. Procedure of translation

The developer of the original English version of the HSS-THRES survey was informed of the objectives of this study and consented to the Japanese translation of the survey (Carol Mancuso, MD, Hospital for Special Surgery, New York, USA, personal communication, 2017). HSS-THRES was translated according to the international guidelines described by Beaton et al. [21] (Fig. 1).

The translation methodology had the following five stages: (1) translation (English to Japanese), (2) synthesis, (3) back translation (Japanese to English), (4) an expert committee review, and (5) pre-testing.

First, two persons independently translated the original English version of the HSS-THRES into Japanese (T1 and T2). Both were Japanese native speakers. Second, the two translations created in the first stage were synthesized by a third party (T-12). Thirdly, two bilingual persons translated the second stage questionnaire back into English (BT1 and BT2). Fourth, an expert committee was organized, in which all translators, health professionals, a methodologist, language experts, and orthopedic experts participated. During this meeting, all versions of the questionnaires (T1, T2, T-12, BT1 and BT2) were combined and a pre-final version of the questionnaire was developed. Finally, the pre-final version was presented to a group of 20 patients [12 women and 8 men; age (standard deviation; SD): 61.7 (8.2) years] to explore the clarity of the questionnaire. All patients were asked whether they understood the items and whether they could interpret the questions correctly. The final Japanese version of the Japanese HSS-THRES was created from these pre-test results (Appendix 1: Available Online). The mean completion time was 11 (SD = 2.7) minutes.

### 2.3. Psychometric scale properties

This study evaluated the psychometric properties of the Japanese version of HSS-THRES according to Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) [22,23], which aims to improve the selection of outcome measurement instruments both in research and clinical practice by developing methodology and practical tools for selecting the most suitable outcome measurement instrument.

#### 2.3.1. Patients

Patients recruited from the Orthopedic Surgery Department of the single participating institution in Japan were enrolled between April 2018 and December 2018. Inclusion criteria were as follows: 1) native Japanese speakers aged at least 20 years and 2) those who could understand and answer the questions. Exclusion criteria were as follows: 1) patients who scheduled revision and 2) those with severe co-morbidities, such as stroke. Sample size was determined according to COSMIN [22,23]. A preoperative questionnaire was used to obtain information about baseline demographic characteristics (e.g., sex, age, living arrangements).

The patients were informed that data from the case would be submitted for publication, and gave their consent. This research has

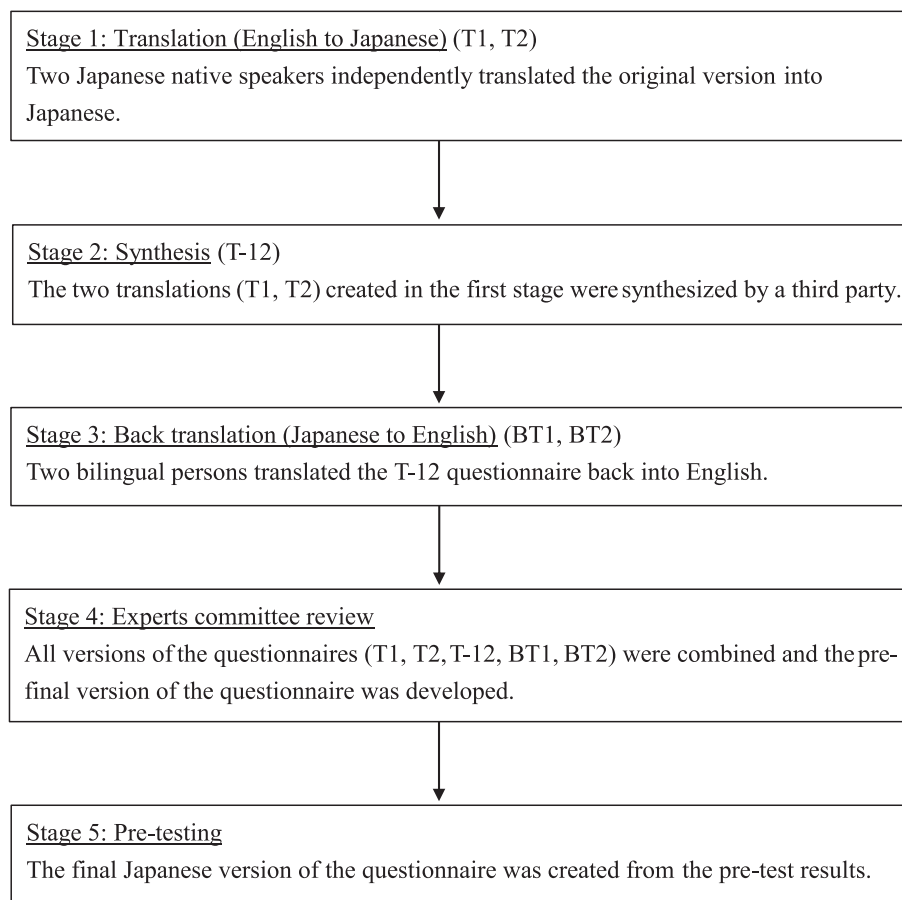


Fig. 1. Procedure of translation.

been approved by the IRB of the authors' affiliated institutions (29–498).

#### 2.3.2. Floor and ceiling effects

Floor and ceiling effects were evaluated to exist when 15% or more of the subjects achieved the lowest (0) or highest (100) points for the total score of HSS-THRES [24].

#### 2.3.3. Internal consistency

Cronbach's alpha was calculated, and its value between 0.70 and 0.95 was taken as an indication of sufficient homogeneity [25].

#### 2.3.4. Test-retest reliability

To determine the test-retest reliability, the surveys were conducted again after a ten-day interval. This time interval between test and retest was chosen because we believe it is long enough to prevent recall of previous answer, though short enough to prevent extensive changes in expectations. The intraclass coefficient correlation (ICC) (two-way random effects model, absolute agreement) was used to test the reliability between the first and second assessment scores and to examine item-by-item agreement [25]. The ICC values were reported with 95% confidence intervals (CI). When the ICC value was above 0.80, the result was considered to have good reproducibility [26].

Additionally, to determine the agreement, Bland and Altman plots were created. In these plots, the mean difference ( $d$ ) between the first and second assessments with corresponding 95% CI and 95% limits of agreement (LOA) were presented ( $d \pm t_{n-1} \times SD_d$ ) [27].

Moreover, Bland and Altman analyses were performed to determine whether any bias occurred.

#### 2.3.5. Construct validity

Construct validity reflects the extent to which a particular measure consistently relates to other measures with theoretically derived hypotheses for the constructs that are being measured. To evaluate the construct validity of the HSS-THRES, we tested nine priori hypotheses, which are about the expected magnitude and direction of relationship with the overall expectation for THA, OHS, and EQ-5D, and the relationship reported in previous research between HSS-THRES and demographic variables. When 75% or more of the hypotheses were confirmed, the construct validity was defined as good [22,25]. The correlation coefficient criterion of judgment was as follows:  $r = 0-0.30$ , low;  $r = 0.31-0.60$ , moderate;  $r = 0.61-1.0$ , high [28].

#### 2.4. Statistical analysis

All statistical analyses were conducted using SPSS version 24.0 (IBM, Chicago) for Windows. Baseline characteristics and questionnaire scores were expressed as mean and SD. For internal consistency, Cronbach  $\alpha$  was used. To determine reliability, ICC and Bland–Altman analysis were performed. In the verification of construct validity, Pearson's correlation coefficient was used, and Mann–Whitney U test was used for comparison between the two groups. P values  $< 0.05$  were considered statistically significant.

### 3. Results

#### 3.1. Patients

Of the 134 patients on the waiting list for THA, 116 patients (86.6%) returned the Japanese HSS-THRES and other questionnaires at baseline. All patients were asked to fill out the second assessment of the Japanese HSS-THRES, and only 109 patients (81.3%) returned the questionnaire. The patients' characteristics and outcome scores on the first and second assessments of the Japanese HSS-THRES and other scales are presented in Table 1. The patients were predominantly female, and the majority of the patients were elderly. The majority of the patients underwent THA for Osteoarthritis.

#### 3.2. Cross-cultural adaption

Given that there are people who used the futon (Japanese traditional bedding) on the floor instead of a bed in Japan, an expert committee discussed about the item "Improve ability to get in or out of a bed, chair, or car". This item asked about the patient's ability to get in or out of a chair and car, apart from a bed, thereby allowing patients who do not use beds to respond accordingly. Thus, we decided to use this item as it is. Although some patients used futon (32.6%), there was no report of difficulty in answering this item in the pre-test.

#### 3.3. Floor and ceiling effects

The Japanese HSS-THRES score of all patients ( $N = 116$ ) ranged from 0 to 100 points. One patient (0.9%) had the lowest score and 13 patients (11.2%) were reported to have the highest score, implying that the Japanese HSS-THRES has no floor or ceiling effects.

#### 3.4. Internal consistency

The Cronbach's alpha coefficient as a measure of internal consistency was 0.90 for the Japanese HSS-THRES, showing a good internal consistency.

#### 3.5. Test-retest reliability

The period of the test-retest reliability was  $12.3 \pm 2.7$  days (range: 7–17 days). The mean score of the second assessment ( $80.8 \pm 18.0$ ) was similar to that of the first test ( $80.6 \pm 18.4$ ). The ICC of the Japanese HSS-THRES score between the first and second assessments was 0.94 (95% CI 0.91–0.96). The ICCs of the individual items ranged from 0.81 (items 1, 7, and 10) to 0.96 (item 12) (Table 2).

To evaluate agreement, the Bland and Altman plot shows that the zero lies within the 95% CI of the mean difference ( $d$ ) between the first and second assessments of the Japanese HSS-THRES, indicating no trend/bias (Fig. 2). The 95% LOA is  $-0.1 \pm 17.7$ .

#### 3.6. Construct validity

Seven of the 9 priori hypotheses for determining the construct validity were verified (77.8%) (Tables 3–5). The following two hypotheses were not confirmed. The hypotheses <the item "physical function" on HSS-THRES and the item "physical function" on OHS are moderately correlated > showed a low correlation with  $r = 0.24$  ( $P < 0.01$ ). The hypothesis <the item Q18 "psychological well-being" on HSS-THRES and the item "anxiety/depression" on EQ-5D are moderately correlated > showed a low correlation with  $r = 0.19$  ( $P = 0.04$ ).

### 4. Discussion

As a result of this validation study, the Japanese HSS-THRES had good cultural acceptability, high reliability, and validity for evaluating patient's expectations before THA. The structure of the original HSS-THRES was not altered, and all items were maintained.

According to the present study, the test-retest reliability of the Japanese HSS-THRES is excellent for the measure of the whole score and each item, with values being higher than the threshold of 0.80 as suggested by Fleiss et al. [26]. Moreover, the Bland and Altman analyses indicated no bias between the first and second measurements. Compared to those in Dutch and French, these results are similar [14,15].

**Table 1**

Patient characteristics and mean scores on the scales for the first and second assessments.

	The first assessment N = 116	The second assessment N = 109
Sex		
Female (n, %)	88 (75.9)	84 (77.1)
Age (mean, SD; range)	62.2 (11.7; 21–87)	62.8 (11.8; 21–87)
Over 65 years old (n, %)	56 (48.3)	55 (50.5)
Diagnosis		
Osteoarthritis (n, %)	98 (84.5)	93 (85.3)
Idiopathic osteonecrosis of femoral head (n, %)	12 (10.3)	11 (10.1)
Rheumatoid arthritis (n, %)	3 (2.6)	2 (1.8)
Traumatic osteonecrosis of femoral head (n, %)	3 (2.6)	3 (2.8)
Comorbidity		
Hypertension (n, %)	36 (31.0)	35 (32.1)
Diabetes mellitus (n, %)	13 (11.2)	13 (11.9)
Heart disease (n, %)	12 (10.3)	12 (11.0)
Employed		
Yes (n, %)	58 (50.0)	56 (51.4)
HSS-THRES score: 0–100		
T1 (mean, SD)	79.9 (18.7)	80.6 (18.4)
T2 (mean, SD)		80.8 (18.0)
Overall expectation for THA: 0–10 (mean, SD)	8.7 (1.5)	8.7 (1.5)
OHS total score: 12–60 (mean, SD)	36.5 (9.1)	36.8 (9.2)
EQ-5D utility score: 0–1 (mean, SD)	0.53 (0.21)	0.53 (0.21)

Abbreviations: SD, standard deviation; HSS-THRES, The Hospital for Special Surgery Hip Replacement Expectations Survey; T1, score from the first assessment of the Japanese version of the Hospital for Special Surgery Hip Replacement Expectations Survey; T2, score from the second assessment of the Japanese version of the Hospital for Special Surgery Hip Replacement Expectations Survey; THA, total hip arthroplasty; OHS, Oxford Hip Score, with higher scores indicating poorer health; EQ-5D utility scores, EuroQol-5 Dimension, with higher scores indicating better health.

**Table 2**

Intraclass correlation coefficients between the first and second assessments of the Japanese HSS-THRES for the total and individual items separately.

	ICC	95% CI
HSS-THRES total score	0.94	0.91–0.96
Q1. Relief of daytime pain	0.81	0.73–0.87
Q2. Relief of pain that interferes with sleep	0.89	0.84–0.93
Q3. Improve ability to walk	0.85	0.77–0.89
Q4. Improve ability to stand	0.88	0.82–0.92
Q5. Get rid of limp	0.93	0.90–0.95
Q6. Remove need for a cane or other assistive device	0.86	0.80–0.90
Q7. Improve ability to climb stairs	0.81	0.72–0.87
Q8. Improve ability to get in or out of a bed, chair or car	0.90	0.85–0.93
Q9. Improve ability to perform daily activities around the home	0.86	0.80–0.90
Q10. Improve ability to perform daily activities away from the home	0.81	0.73–0.87
Q11. Eliminate need for medications	0.96	0.95–0.98
Q12. Be employed for monetary reimbursement	0.96	0.94–0.97
Q13. Improve sexual activity	0.92	0.88–0.94
Q14. Improve ability to exercise or participate in sports	0.92	0.88–0.94
Q15. Improve ability to participate in social activities or recreation	0.88	0.82–0.92
Q16. Improve ability to put on shoes and socks	0.92	0.89–0.95
Q17. Improve ability to cut toenails	0.94	0.91–0.96
Q18. Improve psychological well-being	0.86	0.79–0.90

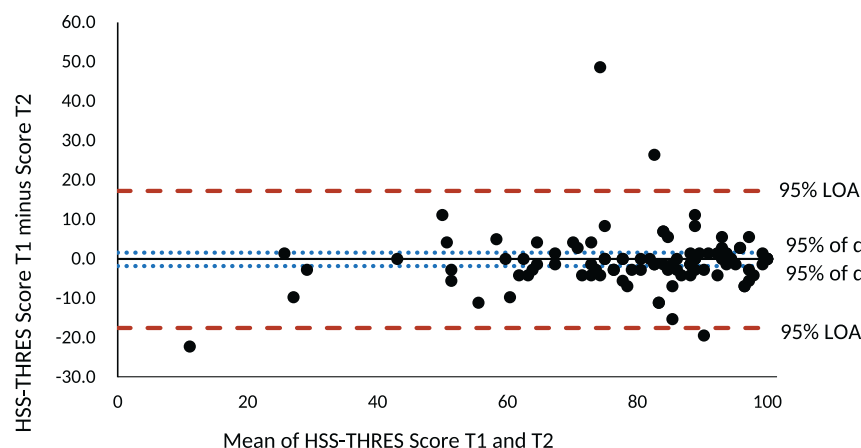
Abbreviations: ICC, intraclass correlation coefficient; CI, confidence interval; HSS-THRES, The Hospital for Special Surgery Hip Replacement Expectations Survey.

The internal consistency of the Japanese HSS-THRES is good (Cronbach's alpha: 0.90). The Cronbach's alpha was slightly higher in the English version of HSS-THRES (Cronbach's alpha: 0.77) and that of previous research than in our Japanese adaptation [9,14–16].

The construct validity of the Japanese HSS-THRES was good, as most of the pre-formulated hypotheses were met. However, 2 of the 9 hypotheses could not be confirmed. The average score of “physical function” on HSS-THRES was  $3.2 \pm 0.7$ , and the average score of Q18 “psychological well-being” on HSS-THRES was  $2.0 \pm 1.1$ , indicating that the patients had great expectations, regardless of the exacerbation of OHS's “physical function” or EQ-5D's “anxiety/depression”. The similar to the finding of the previous

study; the correlation coefficient between HSS-THRES and OHS/SF-12 MCS showed a weak correlation of 0.2 or less [29].

The strength of this study is the evaluation of the psychometric properties based on the methodology proposed by COSMIN [22,23]. Given that there is no instrument available to determine patient expectations that could be considered the gold standard, the Expectation Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) was chosen to determine concurrent validity of the Dutch, French, and Chinese adaptations. However, the psychometric properties of Expectation WOMAC remain unknown. Moreover, the Bland–Altman analysis unfortunately showed a systematic bias [14–16]. Therefore, we deemed that hypothesis



**Fig. 2.** Bland–Altman plot. T1: score from the first assessment of The Hospital for Special Surgery Hip Replacement Expectations Survey (HSS-THRES); T2: score from the second assessment of the HSS-THRES; d: mean difference between the first and second assessments of the survey; LOA: limits of agreement.



**Table 3**

Predefined hypotheses and the confirmation or rejection of the hypotheses.

	N = 116
1. There is a moderate-to-high correlation between the Japanese HSS-THRES scores and overall expectation for THA	Yes
2. The correlation between the Japanese HSS-THRES and OHS is moderate	Yes
3. The correlation between HSS-THRES and disease specific QOL scale (OHS) is higher than that of the general QOL scale (EQ-5D utility scores)	Yes
4. The item "pain" on HSS-THRES and the item "pain" on OHS are moderately correlated	Yes
5. The item "physical function" on HSS-THRES and the item "physical function" on OHS are moderately correlated	No
6. The item Q18 "psychological well-being" on HSS-THRES and the item "anxiety/depression" on EQ-5D are moderately correlated	No
7. The patient's disease duration and diagnosis are unrelated to the Japanese HSS-THRES	Yes
8. The patient's age is unrelated to the Japanese HSS-THRES	Yes
9. The patient's sex is unrelated to the Japanese HSS-THRES	Yes
	7/9 (77.8%)

Abbreviations: HSS-THRES, The Hospital for Special Surgery Hip Replacement Expectations Survey; OHS, Oxford Hip Score; EQ-5D, EuroQol-5 Dimension.

**Table 4**

Correlation data for HSS-THRES at the first assessment.

	Correlation with HSS-THRES r (P-value)	Correlation with HSS-THRES pain items r (P-value)	Correlation with HSS-THRES physical function items r (P-value)	N = 116 Correlation with HSS-THRES Q18 psychological well-being item r (P-value)
Overall expectation for THA	0.67 (<0.01)	—	—	—
OHS	0.37 (<0.01)	—	—	—
OHS pain items	—	0.35 (<0.01)	—	—
OHS physical function items	—	—	0.24 (<0.01)	—
EQ-5D utility scores	-0.20 (0.03)	—	—	—
EQ-5D anxiety/depression items	—	—	—	0.19 (0.04)
Disease duration	-0.09 (0.32)	—	—	—

Pearson's correlation coefficient.

Abbreviations: HSS-THRES, The Hospital for Special Surgery Hip Replacement Expectations Survey; THA, Total hip arthroplasty; OHS, Oxford Hip Score, with higher scores indicating poorer health; EQ-5D, Euro-QOL 5 Dimension; EQ-5D utility scores, with higher scores indicating better health; EQ-5D anxiety/depression, items, with higher scores indicating poorer health.

testing to determine the construct validity was appropriate as a methodology. Furthermore, in the cross-cultural adaptation of the Japanese version of HSS-THRES, all patients completed the questionnaires without any difficulties, indicating that the translated questionnaires had good cultural acceptability.

There were some limitations in this study that should be considered. This study was limited to one institution, which may not be representative of the whole Japanese population. However, this study was similar to the demographic variables from a previous study conducted in Japan [30].

Our study findings demonstrated that it is possible to grasp the Japanese patients' expectation to avoid making unrealistic expectations from them. Health providers recommend early post-operative treatment and rehabilitation according to patient's individual expectations by using the Japanese version of HSS-

THRES. Furthermore, HSS-THRES has Dutch, French, Chinese, and German versions [13–16]. The Japanese version developed by this research makes it possible to compare expectations of patients from different countries and cultures.

## 5. Conclusions

This study translated the English version of HSS-THRES into Japanese according to the international guidelines described by Beaton et al. [21] and evaluated its psychometric properties based on the methodology proposed by COSMIN [22,23]. The Japanese version of the HSS-THRES has good internal consistency, reliability, and construct validity, and no floor and ceiling effects. Therefore, we recommend its use as a tool for evaluating patients' expectations before THA.

## Conflict of interest

None.

## Funding

This work was supported by the Grants-in-Aid for Scientific Research of Japan Society for the Promotion of Science, Grant Number Research (C) FAG5K11756.

## Acknowledgments

The authors are grateful to the THA patients and the healthcare providers at Kyushu University Hospital for their cooperation.

**Table 5**

Significant difference in HSS-THRES scores at the first assessments.

	HSS-THRES score	N = 116 P-value
Diagnosis		
Osteoarthritis	81.5 ± 15.5	0.47
Others	71.9 ± 29.6	
Age		
Younger (≤62 years old)	80.9 ± 18.3	0.57
Older (>62 years old)	79.1 ± 19.2	
Sex		
Female	82.0 ± 15.5	0.23
Male	73.4 ± 25.8	

Mann–Whitney U test.

Abbreviations: HSS-THRES, The Hospital for Special Surgery Hip Replacement Expectations Survey.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jos.2019.07.006>.

## References

- [1] Stratford PW, Kennedy DM. Performance measures were necessary to obtain a complete picture of osteoarthritic patients. *J Clin Epidemiol* 2006 Feb;59(2):160–7.
- [2] Mizner RL, Petterson SC, Clements KE, Zeni Jr JA, Irrgang JJ, Snyder-Mackler L. Measuring functional improvement after total knee arthroplasty requires both performance-based and patient-report assessments: a longitudinal analysis of outcomes. *J Arthroplast* 2011 Aug;26(5):728–37.
- [3] Tekin B, Unver B, Karatosun V. Expectations in patients with total knee arthroplasty. *Acta Orthop Traumatol Turcica* 2012;46(3):174–80.
- [4] Mancuso CA, Salvati EA, Johanson NA, Peterson MG, Charlson ME. Patients' expectations and satisfaction with total hip arthroplasty. *J Arthroplast* 1997 Jun;12(4):387–96.
- [5] Gonzalez Saenz de Tejada M, Escobar A, Bilbao A, Herrera-Espineira C, Garcia-Perez L, Aizpuru F, Sarasqueta C. A prospective study of the association of patient expectations with changes in health-related quality of life outcomes, following total joint replacement. *BMC Musculoskelet Disord* 2014 Jul 23;15:248.
- [6] Mancuso CA, Jout J, Salvati EA, Sculco TP. Fulfillment of patients' expectations for total hip arthroplasty. *J Bone Joint Surg Am* 2009 Sep;91(9):2073–8.
- [7] Tilbury C, Haanstra TM, Leichtenberg CS, Verdegaal SH, Ostelo RW, de Vet HC, Nelissen RG, Vliet Vlieland TP. Unfulfilled expectations after total hip and knee arthroplasty surgery: there is a need for better preoperative patient information and education. *J Arthroplast* 2016 Oct;31(10):2139–45.
- [8] Iversen MD, Daltroy LH, Fossel AH, Katz JN. The prognostic importance of patient pre-operative expectations of surgery for lumbar spinal stenosis. *Patient Educ Couns* 1998 Jun;34(2):169–78.
- [9] Mancuso CA, Graziano S, Briskie LM, Peterson MG, Pellicci PM, Salvati EA, Sculco TP. Randomized trials to modify patients' preoperative expectations of hip and knee arthroplasties. *Clin Orthop Relat Res* 2008 Feb;466(2):424–31.
- [10] Vissers MM, Bussmann JB, Verhaar JA, Busschbach JJ, Bierma-Zeinstra SM, Reijman M. Psychological factors affecting the outcome of total hip and knee arthroplasty: a systematic review. *Semin Arthritis Rheum* 2012 Feb;41(4):576–88.
- [11] Haanstra TM, van den Berg T, Ostelo RW, Poolman RW, Jansma EP, Cuijpers P, de Vet HC. Systematic review: do patient expectations influence treatment outcomes in total knee and total hip arthroplasty? *Health Qual Life Outcomes* 2012 Dec;10:152.
- [12] Mancuso CA, Sculco TP, Salvati EA. Patients with poor preoperative functional status have high expectations of total hip arthroplasty. *J Arthroplast* 2003 Oct;18(7):872–8.
- [13] Balck F, Kirschner S, Jeszenszky C, Lippmann M, Gunther KP. [Validity and reliability of the German version of the HSS expectation questionnaire on hip joint replacement]. *Z für Orthop Unfallchirurgie* 2016 Dec;154(6):606–11 [in German].
- [14] Neuprez A, Delcour JP, Fatemi F, Gillet P, Mawet M, Francois G, Bruyère O, Crielaard JM, Gosset C, Reginster JY. Development and validation of the French version of a tool assessing patient's expectations in lower limb osteoarthritis. *J Orthop* 2014 Jun 21;12(1):46–57.
- [15] van den Akker-Scheek I, van Raay JJ, Reininga IH, Bulstra SK, Zijlstra W, Stevens M. Reliability and concurrent validity of the Dutch hip and knee replacement expectations surveys. *BMC Musculoskelet Disord* 2010 Oct 19;11:242.
- [16] Wang C, Zhang C, Liu DL, Tong WW, He CR, Huang X, Xu WD. Simplified Chinese version of hip and knee replacement expectations surveys in patients with osteoarthritis and ankylosing spondylitis: cross-cultural adaptation, validation and reliability. *BMC Musculoskelet Disord* 2018 Jul 21;19(1):247.
- [17] Engel C, Hamilton NA, Potter PT, Zautra AJ. Impact of two types of expectancy on recovery from total knee replacement surgery (TKR) in adults with osteoarthritis. *Behav Med* 2004 Fall;30(3):113–23.
- [18] Harris KK, Price AJ, Beard DJ, Fitzpatrick R, Jenkinson C, Dawson J. Can pain and function be distinguished in the Oxford Hip Score in a meaningful way? : an exploratory and confirmatory factor analysis. *Bone Joint Res* 2014 Nov;3(11):305–9.
- [19] Uesugi Y, Makimoto K, Fujita K, Nishii T, Sakai T, Sugano N. Validity and responsiveness of the Oxford hip score in a prospective study with Japanese total hip arthroplasty patients. *J Orthop Sci* 2009 Jan;14(1):35–9.
- [20] Wolfe F, Hawley DJ. Measurement of the quality of life in rheumatic disorders using the EuroQol. *Br J Rheumatol* 1997 Jul;36(7):786–93.
- [21] Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)* 2000 Dec;25(24):3186–91.
- [22] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HCW. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010 May;19(4):539–49.
- [23] Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, Terwee CB. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018 May;27(5):1171–9.
- [24] McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 1995 Aug;4(4):293–307.
- [25] Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007 Jan;60(1):34–42.
- [26] Fleiss JL, Shrout PE. The effects of measurement errors on some multivariate procedures. *Am J Public Health* 1977 Dec;67(12):1188–91.
- [27] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986 Feb;1(8476):307–10.
- [28] Andresen EM. Criteria for assessing the tools of disability outcomes research. *Arch Phys Med Rehabil* 2000 Dec;81(12 Suppl 2):S15–20.
- [29] Scott CE, Bugler KE, Clement ND, MacDonald D, Howie CR, Biant LC. Patient expectations of arthroplasty of the hip and knee. *J Bone Joint Surg Br* 2012 Jul;94(7):974–81.
- [30] Fujita K, Makimoto K, Mawatari M. Three-year follow-up study of health related QOL and lifestyle indicators for Japanese patients after total hip arthroplasty. *J Orthop Sci* 2016 Mar;21(2):191–8.