

Validity and Reliability of the Thai Version of the International Hip Outcome Tool in Young to Middle-Aged Physically Active Adults With Hip Disorders

Pisit Lertwanich,^{*†} MD, Siwa Loog-in,[†] MD, Chaturong Pornrattanamaneewong,[†] MD, and Phob Ganokroj,[†] MD

Investigation performed at the Mahidol University, Bangkok, Thailand

Background: The International Hip Outcome Tool (iHOT-33), which is a self-administered questionnaire that contains 33 questions, was developed to evaluate young to middle-aged physically active adults with hip pathology.

Purpose: To translate and cross-culturally adapt the iHOT-33 to create a Thai version of the iHOT-33 (TH-iHOT-33) and to determine the validity and reliability of the TH-iHOT-33.

Study Design: Cohort study (diagnosis); Level of evidence, 3.

Methods: Translation and cross-cultural adaptation of the iHOT-33 were performed according to international guidelines. Patients aged 18 to 60 years with hip disorders who had a score ≥ 4 on a modified version of the Tegner activity scale completed the TH-iHOT-33 twice, 2 weeks apart. The Thai versions of the HOOS (Hip disability and Osteoarthritis Outcome Score) and EQ-5D-5L (European Quality of Life–5 Dimensions–5 Levels) were also completed at baseline. At 2 weeks, the Global Perceived Effect scale was completed in addition to the TH-iHOT-33. The statistical testing for validity included construct validity and hypotheses testing. Internal consistency, test-retest reliability, and measurement error were analyzed to determine reliability.

Results: The iHOT-33 was successfully translated and cross-culturally adapted to create the TH-iHOT-33. Seventy-two patients with various hip disorders participated in this study. Of the predefined hypotheses used for construct validity, 86% were confirmed. A Cronbach alpha of 0.98 (95% CI, 0.97-0.99) demonstrated excellent internal consistency. Test-retest reliability was high, with an intraclass correlation coefficient of 0.83 (95% CI, 0.73-0.90). The standard error of measurement was 5.1 points. The smallest detectable change was 14.2 points at the individual level and 1.9 points at the group level. The minimal important change was 10.9 points. The TH-iHOT-33 had no observable floor or ceiling effects.

Conclusion: The TH-iHOT-33 is a valid and reliable instrument for assessing the functional status and health-related quality of life of young to middle-aged physically active Thai adults with hip disorders.

Keywords: validity; reliability; cross-cultural adaptation; International Hip Outcome Tool; iHOT-33; hip disorders

Patient-reported outcomes (PROs) are important tools for evaluation of the patient's health condition from the patient's perspective.¹⁴ Either generic or disease-specific questionnaires have been used in research and clinical practice for assessing a change in a patient's health status after an intervention.^{3,14} PROs are considered the gold standard for measuring outcomes for nonsurgical and surgical treatment.³¹ Traditional PROs for patients with hip disorders were usually developed for patients with osteoarthritis or for patients who underwent hip replacement surgery.¹⁹ Over the past decades, various conditions that cause

hip pain in young adults have gained recognition, especially femoroacetabular impingement syndrome and labral tears.²⁷ Several hip-specific PROs were developed for assessing the functional status and health-related quality of life (QoL) of young to middle-aged patients who were physically active in their activities of daily living, occupations, and sports.^{4,6,13,15,30} Kemp et al¹¹ showed that the modified Harris Hip Score and the subscales related to activities of daily living of the Hip disability and Osteoarthritis Outcome Score (HOOS), Hip Outcome Score, and HAGOS (Copenhagen Hip and Groin Outcome Score) demonstrated ceiling effects when they were used to evaluate outcomes of hip arthroscopic surgery.

The International Hip Outcome Tool (iHOT-33) is a self-administered questionnaire that was developed by the

The Orthopaedic Journal of Sports Medicine, 9(9), 23259671211042017

DOI: 10.1177/23259671211042017

© The Author(s) 2021

This open-access article is published and distributed under the Creative Commons Attribution - NonCommercial - No Derivatives License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits the noncommercial use, distribution, and reproduction of the article in any medium, provided the original author and source are credited. You may not alter, transform, or build upon this article without the permission of the Author(s). For article reuse guidelines, please visit SAGE's website at <http://www.sagepub.com/journals-permissions>.

Multicenter Arthroscopy of the Hip Outcomes Research Network to evaluate young to middle-aged active patients with symptomatic hip pathology.¹⁵ This QoL PRO measure has good psychometric properties relating to validity, reliability, and responsiveness.^{11,15,25,31} The iHOT-33 tool has been translated and validated in several languages.^{1,12,23,26,32} To use the iHOT-33 to evaluate Thai patients with hip pathology, the aim of this study was to translate and cross-culturally adapt the iHOT-33 questionnaire to create the Thai version of the iHOT-33 (TH-iHOT-33) and to determine the validity and reliability of the TH-iHOT-33. We hypothesized that the TH-iHOT-33 would be a valid and reliable instrument for assessing the functional status and health-related QoL of young to middle-aged physically active Thai adults with hip disorders.

METHODS

This cross-sectional study was conducted from August 2018 to September 2020. The study consisted of 2 parts: (1) translation and cross-cultural adaptation of the iHOT-33 from the English-language version into the Thai-language version and (2) evaluation of the validity and reliability of the TH-iHOT-33 in young to middle-aged physically active Thai adults with hip disorders. The protocol for this study was approved by an institutional review board, and informed consent was obtained from all study participants.

Translation and Cross-Cultural Adaptation

After permission was received from the original authors,¹⁵ the iHOT-33 was translated into the Thai language following the international guidelines for the process of cross-cultural adaptation of self-report measures.² Forward translation was independently conducted by 2 bilingual translators: an orthopaedic surgeon and an uninformed professional translator. Both translations were compared and synthesized into a preliminary TH-iHOT-33 questionnaire in a consensus meeting. Two professional translators then independently translated that Thai version back into English. Both translators were unaware of the study objectives and the original questionnaire. These back-translations were then compared with the original version by an expert committee. Minor discrepancies regarding word choice were discussed and resolved by consensus; this resulted in a prefinal TH-iHOT-33. This prefinal version was tested on 20 Thai patients with hip disorders to assess understanding and to identify problems. Patient understanding was good and no problems were identified, so the prefinal version was adopted as the final version of TH-iHOT-33 with no additional changes.

Study Procedure

The procedure to evaluate the validity and reliability of the TH-iHOT-33 was performed according to the recommendations in the COSMIN checklist (Consensus-Based Standards for the Selection of Health Measurement Instruments).^{16,17} Patients with hip disorders were included if they were aged 18 to 60 years and they had a score ≥ 4 on a modified version of the Tegner activity scale.²⁸ This activity level was used to identify physically active individuals in the development of the original iHOT-33.¹⁵ Patients were excluded if they had insufficient understanding of the Thai language or cognitive impairment that prevented reliable completion of the questionnaires used in this study. Patients who had disorder of the back or the contralateral lower extremity, active hip joint infection, or musculoskeletal malignancy were also excluded.

All patients were first evaluated in an outpatient setting. They were asked to complete 3 paper-based questionnaires: the TH-iHOT-33, the Thai version of the HOOS,³³ and the Thai version of the European Quality of Life–5 Dimensions–5 Levels (EQ-5D-5L).²² They were also provided with a stamped preaddressed envelope and the second set of questionnaires, including the TH-iHOT-33 questionnaire and the Global Perceived Effect (GPE) scale, which patients were instructed to complete 2 weeks after the first set of questionnaires were completed. This interval was considered sufficiently long to prevent recall of previous responses and recent enough to ensure minimal clinical changes.²⁹ All PROs used in this study were paper-based questionnaires.

Questionnaires

The iHOT-33 is a disease-specific self-report questionnaire that consists of 33 questions. These questions are divided into 4 domains: symptoms and functional limitations (16 items); sports and recreational activities (6 items); job-related concerns (4 items); and social, emotional, and lifestyle concerns (7 items).¹⁵ Every question uses a visual analog scale response format that is designed for conversion to computer- and web-based self-administration. Each question is scored from 0 to 100 points. The total score represents the sum of all answered questions divided by the number of answered questions. Higher scores represent better physical function and health-related QoL.¹⁵ It is possible to calculate a separate score for each domain, but there was no attempt to score the 4 domains separately.¹⁵

The HOOS is a disease-specific questionnaire to evaluate patients with hip disability with or without osteoarthritis.^{19,20} This questionnaire has been translated into Thai and has shown good psychometric properties.³³ The HOOS consists

*Address correspondence to Pisit Lertwanich, MD, Department of Orthopaedic Surgery, Faculty of Medicine Siriraj Hospital, Mahidol University, 2 Wanglang Road, Bangkoknoi, Bangkok 10700, Thailand (email: pisit.ler@mahidol.ac.th).

[†]Department of Orthopaedic Surgery, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok, Thailand.

Final revision submitted April 21, 2021; accepted June 8, 2021.

One or more of the authors has declared the following potential conflict of interest or source of funding: This study was funded by a research grant (R016132035) from the Faculty of Medicine, Siriraj Hospital, Mahidol University. AOSM checks author disclosures against the Open Payments Database (OPD). AOSM has not conducted an independent investigation on the OPD and disclaims any liability or responsibility relating thereto.

Ethical approval for this study was obtained from Siriraj Hospital Institutional Review Board, No. 339/2561(EC4).

of 40 items with a 5-point Likert scale format. These 40 items are divided into 5 subscales: symptoms, pain, activities of daily living, sport and recreation function, and hip-related QoL. The HOOS can be transformed into a 0 to 100 scale of total score or separate subscale.²⁰ Higher scores represent better function and fewer symptoms.

The EQ-5D is a generic health-related QoL assessment tool.^{7,24} This questionnaire was developed by the EuroQol Group.²⁴ It consists of 5 items assessing 5 dimensions of health status: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The EQ-5D-5L is a new version of the questionnaire, and each item has 5 levels of response.⁷ The calculated utility score of the EQ-5D-5L ranges from 0 to 1. Higher scores indicate better health-related QoL. This questionnaire has been translated into many languages worldwide, including Thai.²²

The GPE scale is used to assess changes in the patient's health status according to the patient's perception. The scale asks the patient to rate how much his or her condition has deteriorated or improved from baseline.¹⁰ This study used an 11-point GPE scale ranging from -5 (vastly worse) to 0 (no change) to +5 (completely recovered).⁵ In the second assessment, the GPE scale was completed with the TH-iHOT-33 to identify a group of patients with no clinically relevant change (who had a GPE scale score of -1, 0, or +1) for assessment of test-retest reliability.

Statistical Analysis

Statistical analyses were performed using PASW Statistics for Windows (Version 18.0; SPSS Inc). A *P* value <.05 was considered statistically significant. A Bland-Altman plot was also produced using MedCalc Statistical Software (Version 16.4.3). Patient demographic and clinical characteristics are summarized using descriptive statistics. Categorical variables are presented as number and percentage, and normally distributed continuous variables are given as mean, standard deviation, and range.

Validity

Construct validity represents the degree to which the scores of a PRO measure are consistent with the scores of another relevant PRO measure.^{17,25} The construct validity of the TH-iHOT-33 was examined by testing a priori hypotheses specific to expected associations between the TH-iHOT-33 and the HOOS and EQ-5D-5L (Table 1). The Spearman rank correlation coefficient (r_s) was used to assess these associations (0-0.3, negligible correlation; 0.3-0.5, low; 0.5-0.7, moderate; 0.7-0.9, high; 0.9-1.0, very high).¹⁸

Reliability

Reliability is defined as the degree to which a PRO measure is free from measurement error.¹⁷ Reliability of the TH-iHOT-33 was evaluated in 3 measurement properties: internal consistency, test-retest reliability, and measurement error.¹⁷

TABLE 1
A Priori Hypotheses Concerning the Correlations
Between the Thai Version of the iHOT-33
and the HOOS and the EQ-5D-5L^a

Hypothesis	r_s (95% CI) ^b	Confirmed (Yes/No)
A correlation of at least 0.7 and in the same direction between the TH-iHOT-33 score and the HOOS total score	0.83 (0.77-0.91)	Yes
A correlation of at least 0.7 and in the same direction between the TH-iHOT-33 score and the HOOS symptoms subscale	0.72 (0.55-0.86)	Yes
A correlation of at least 0.7 and in the same direction between the TH-iHOT-33 score and the HOOS pain subscale	0.72 (0.61-0.84)	Yes
A correlation of at least 0.5 and in the same direction between the TH-iHOT-33 score and the HOOS ADL subscale	0.67 (0.56-0.81)	Yes
A correlation of at least 0.7 and in the same direction between the TH-iHOT-33 score and the HOOS sport and recreation subscale	0.67 (0.59-0.85)	No
A correlation of at least 0.5 and in the same direction between the TH-iHOT-33 score and the HOOS health-related QoL subscale	0.67 (0.52-0.81)	Yes
A correlation of at least 0.5 and in the same direction between the TH-iHOT-33 score and the EQ-5D-5L score	0.73 (0.58-0.85)	Yes

^aADL, activities of daily living; EQ-5D-5L, European Quality of Life-5 Dimensions-5 Levels; HOOS, Hip disability and Osteoarthritis Outcome Score; iHOT-33, International Hip Outcome Tool; QoL, quality of life; TH-iHOT-33, Thai version of the International Hip Outcome Tool.

^bValues are given as Spearman rank correlation coefficients.

Internal consistency represents the degree of interrelatedness among the items in the same domain or questionnaire, thus measuring the same concept.^{17,29} The homogeneity of questions was measured by Cronbach alpha coefficient.^{16,17} A Cronbach alpha value within the range of 0.70 to 0.95 is considered ideal.²⁹ A value <0.70 indicates a lack of correlation among items, and a value >0.95 indicates redundancy of questions.^{25,29}

Test-retest reliability was calculated in a group of patients who had a GPE scale score of -1, 0, or +1 in the second response, which was completed 2 weeks after the first set of questionnaires.^{1,26,29} Intraclass correlation coefficient (ICC; 2-way random-effects model) was used to measure test-retest reliability. An ICC value ≥ 0.70 indicates high reliability in a sample size of at least 50 patients.²⁹

Measurement error was evaluated by the standard error of measurement (SEM), which was converted to the smallest detectable change (SDC).¹⁶ The SEM was calculated

TABLE 2
Demographic and Clinical Characteristics (N = 72)^a

Characteristics	No. (%) or Mean ± SD (Range)
Sex	
Male	35 (48.6)
Female	37 (51.4)
Age, y	39.4 ± 11.1 (18-59)
BMI, kg/m ²	23.0 ± 3.8 (15.6-37.2)
Tegner activity level	
4	67 (93.1)
≥5	5 (6.9)
Diagnosis	
FAI syndrome	13 (18.1)
Labral tear	12 (16.7)
Osteonecrosis	10 (13.9)
Hip dysplasia	8 (11.1)
Capsulitis	7 (9.7)
Synovial chondromatosis	7 (9.7)
Ligamentum teres tear	4 (5.6)
Other	11 (15.3)

^aBMI, body mass index; FAI, femoroacetabular impingement.

using the formula $SD \times \sqrt{1 - ICC}$,³⁴ and the standard deviation (SD) from all patients at the initial assessment was applied. The SEM was converted to the SDC at the individual level (SDC_{ind}) with the formula $SEM \times 1.96 \times \sqrt{2}$.³⁴ The SDC_{ind} indicates the smallest within-person change in score that can be interpreted as a real change above measurement error.²⁹ The SEM was also converted to the SDC at the group level (SDC_{group}) by dividing the SDC_{ind} by \sqrt{n} .³⁴ The minimal important change (MIC) was calculated by a distribution-based method using the formula $0.5 \times SD$, which was proposed by Norman et al,²¹ and the SD from all patients at the initial assessment was used. Bland-Altman analysis was used to evaluate agreement between the first and second administration of the TH-iHOT-33.

Floor and Ceiling Effects

Floor or ceiling effects were considered significant if >15% of patients had the lowest or highest possible scores, respectively.²⁹ These effects indicate limitations of the questionnaire in terms of its content validity, reliability, and responsiveness.²⁹

RESULTS

Translation and Cross-Cultural Adaptation

The iHOT-33 was successfully translated from the original version into the Thai version. During the translation process, none of the items or responses in the questionnaire required cultural adaptation.

Patient Characteristics and Generalizability

From August 2018 to September 2020, a total of 72 patients with hip disorders completed the TH-iHOT-33

TABLE 3
Distribution of Scores, Floor Effects, and Ceiling Effects of the TH-iHOT-33 (N = 72)^a

TH-iHOT-33 Item	Score, Mean ± SD	Floor Effect, No. (%)	Ceiling Effect, No. (%)	MIC
Q1	46.9 ± 29.2	1 (1.4)	1 (1.4)	14.6
Q2	53.7 ± 28.2	0 (0.0)	2 (2.8)	14.1
Q3	53.9 ± 30.4	2 (2.8)	3 (4.2)	15.2
Q4	58.0 ± 28.5	2 (2.8)	6 (8.3)	14.3
Q5	55.4 ± 29.9	1 (1.4)	2 (2.8)	14.9
Q6	45.5 ± 28.8	1 (1.4)	2 (2.8)	14.4
Q7	60.5 ± 28.4	0 (0.0)	3 (4.2)	14.2
Q8	54.2 ± 34.0	4 (5.6)	6 (8.3)	17.0
Q9	57.7 ± 28.2	0 (0.0)	1 (1.4)	14.1
Q10	56.0 ± 28.6	0 (0.0)	4 (5.6)	14.3
Q11	57.7 ± 29.3	0 (0.0)	3 (4.2)	14.6
Q12	60.5 ± 28.7	0 (0.0)	5 (6.9)	14.3
Q13	55.0 ± 29.1	0 (0.0)	1 (1.4)	14.5
Q14	54.5 ± 30.7	1 (1.4)	2 (2.8)	15.3
Q15	59.4 ± 32.4	2 (2.8)	3 (4.2)	16.2
Q16	48.0 ± 28.5	0 (0.0)	0 (0.0)	14.3
Q17	46.5 ± 30.8	3 (4.2)	3 (4.2)	15.4
Q18	49.7 ± 27.9	1 (1.4)	2 (2.8)	13.9
Q19	39.4 ± 30.4	8 (11.1)	1 (1.4)	15.2
Q20	43.1 ± 31.6	4 (5.6)	4 (5.6)	15.8
Q21	45.4 ± 31.4	2 (2.8)	2 (2.8)	15.7
Q22	35.9 ± 29.7	4 (5.6)	3 (4.2)	14.8
Q23	49.5 ± 29.4	1 (1.4)	1 (1.4)	14.7
Q24	39.4 ± 31.0	2 (2.8)	2 (2.8)	15.5
Q25	42.8 ± 31.9	3 (4.2)	2 (2.8)	16.0
Q26	48.1 ± 31.8	3 (4.2)	1 (1.4)	15.9
Q27	43.4 ± 34.2	4 (5.6)	4 (5.6)	17.1
Q28	65.2 ± 28.1	0 (0.0)	3 (4.2)	14.1
Q29	47.6 ± 30.6	4 (5.6)	3 (4.2)	15.3
Q30	54.7 ± 29.5	1 (1.4)	2 (2.8)	14.8
Q31	57.0 ± 32.1	2 (2.8)	6 (8.3)	16.0
Q32	65.9 ± 28.4	1 (1.4)	2 (2.8)	14.2
Q33	43.7 ± 31.5	4 (5.6)	2 (2.8)	15.8
Subscale				
Symptoms	54.8 ± 22.1	0 (0.0)	0 (0.0)	11.0
Sports	42.9 ± 26.1	0 (0.0)	1 (1.4)	13.0
Job	44.9 ± 27.6	0 (0.0)	1 (1.4)	13.8
Social	52.0 ± 26.2	0 (0.0)	0 (0.0)	13.1
Total score	51.1 ± 21.7	0 (0.0)	0 (0.0)	10.9

^aMIC, minimal important change; Q, question; TH-iHOT-33, Thai version of the International Hip Outcome Tool.

questionnaire twice, 2 weeks apart. The demographic and clinical characteristics of study patients are shown in Table 2. In the first administration, the TH-iHOT-33 scores ranged from 13.6 to 95.3. The distribution of scores is shown in Table 3. In the second administration, 53 patients (73.6%) had a GPE scale ranging from -1 to +1. The TH-iHOT-33 scores of this group were used to assess test-retest reliability. Concerning the GPE scale, 3 patients (4.2%) scored less than -1, and 16 patients (22.2%) scored higher than +1.

Some items of the iHOT-33 contained the option "I do not do this" or "This is not relevant to me." These answers were considered missing values that could be omitted, and the

TABLE 4
Test-Retest Reliability of the TH-iHOT-33 (n = 53)^a

TH-iHOT-33	Test 1, Mean ± SD	Test 2, Mean ± SD	Mean Difference (95% CI)	SEM	SDC _{ind}	SDC _{grp}	ICC (95% CI)
Subscale							
Symptoms	52.3 ± 22.0	48.3 ± 22.1	4.0 (0.3 to 7.7)	5.8	16.0	2.2	0.82 (0.70-0.89)
Sports	41.7 ± 26.3	39.8 ± 23.9	1.9 (-2.2 to 6.0)	6.3	17.6	2.4	0.82 (0.71-0.89)
Job (n = 45)	43.2 ± 27.8	39.8 ± 23.3	3.5 (-2.6 to 9.5)	11.3	31.3	4.7	0.69 (0.50-0.82)
Social	49.9 ± 26.7	49.6 ± 25.6	0.3 (-3.6 to 4.2)	5.4	15.1	2.1	0.85 (0.76-0.91)
Total score	49.1 ± 21.9	46.3 ± 21.1	2.8 (-0.7 to 6.2)	5.1	14.2	1.9	0.83 (0.73-0.90)

^aICC, intraclass correlation coefficient; SDC_{grp}, smallest detectable change at the group level; SDC_{ind}, smallest detectable change at the individual level; SEM, standard error of measurement; TH-iHOT-33, Thai version of the International Hip Outcome Tool.

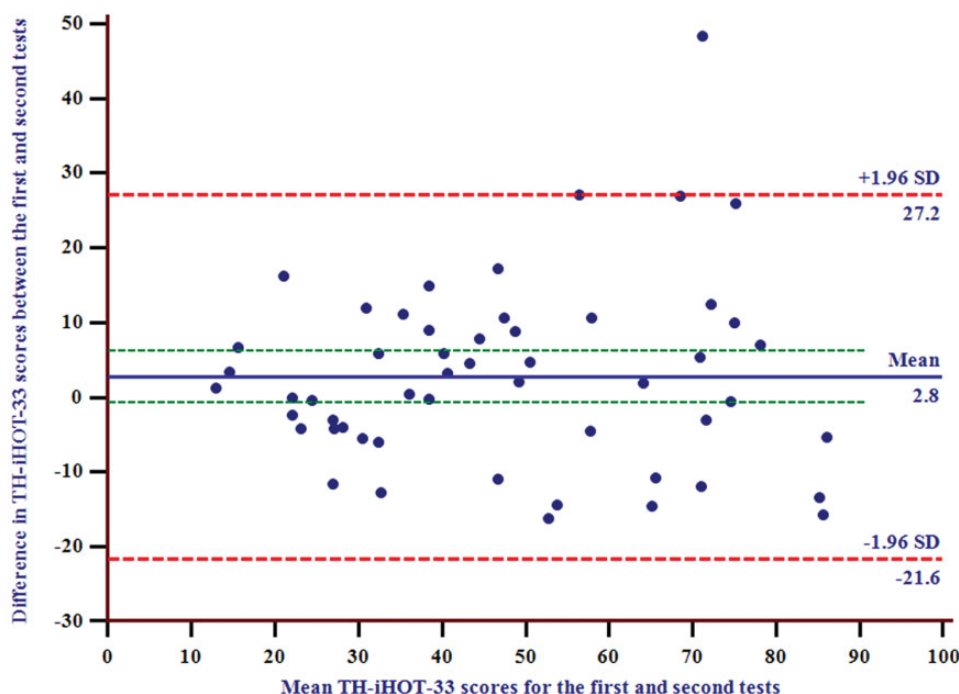


Figure 1. Bland-Altman plot to evaluate the test-retest reliability of the Thai version of the International Hip Outcome Tool (TH-iHOT-33). The dashed green lines indicate 95% CI of mean of difference.

total score was calculated from the remaining answers.¹⁵ Missing values, as defined by this reason, were found in 313 (6.6%) of 4752 items. Twelve patients (16.7%) reported that they did not work, so they skipped the job-related concerns section. Two of those patients did not work as a result of their hip pathology, and the other 10 cited reasons other than hip pathology for why they were unemployed. Eight patients in this group were excluded from assessment of test-retest reliability of the job-related concerns subscale.

Validity

The Spearman rank correlation coefficients among the TH-iHOT-33, the Thai version of the HOOS, and the Thai version of the EQ-5D-5L are presented in Table 1. The TH-

iHOT-33 had strong correlation with the HOOS total score and the EQ-5D-5L. Seven of the a priori hypotheses were tested, and 6 were confirmed (86.0%).

Reliability

A Cronbach alpha of 0.98 (95% CI, 0.97-0.99) demonstrated excellent internal consistency for the TH-iHOT-33. The Cronbach alpha for each subscale was also excellent, with a range from 0.91 to 0.95. The results of test-retest reliability measures are shown in Table 4. The ICC value for the TH-iHOT-33 was 0.83 (95% CI, 0.73-0.90), which indicated high test-retest reliability. The SEM was 5.1 points, and the SDCs at the individual and group levels were 14.2 and 1.9 points, respectively. The MIC of the TH-iHOT-33 was 10.9 points

(Table 3). A Bland-Altman plot of the difference between the first and second administration of the TH-iHOT-33 is shown in Figure 1.

Floor and Ceiling Effects

There were no patients who scored the lowest or highest possible score of the TH-iHOT-33. Therefore, this questionnaire demonstrated no floor or ceiling effects. The floor and ceiling effects of each item and subscale were also evaluated (Table 3). Among the 33 questions, some floor and ceiling effects were identified, but none were seen in >15% of the participants.

DISCUSSION

In this study, the TH-iHOT-33 was developed and the psychometric properties were assessed. The questionnaire was successfully translated into the Thai language using the international guidelines of the process of cross-cultural adaptation of self-report measures. The psychometric properties showed that the TH-iHOT-33 is a valid, internally consistent, and reliable measurement tool for evaluating young to middle-aged physically active Thai patients with hip disorders.

The iHOT-33 is a PRO measure that was developed for assessing the functional status and health-related QoL of young to middle-aged physically active patients with hip pathology.¹⁵ The iHOT-33 shows good content validity because orthopaedic surgeons, physical therapists, and patients were involved in item generation during the initial phase of development.^{15,31} In a systematic review that evaluated the clinimetric properties of 9 questionnaires for patients with hip and groin disability, the iHOT-33 was one of the recommended PROs for young to middle-aged adults with hip joint-related pain who were undergoing nonsurgical treatment or hip arthroscopy.³¹ The International Hip-Related Pain Research Network recently stated that the HAGOS and iHOT questionnaires, the extended and short versions, are the most appropriate PRO measures in young and middle-aged active adults with hip-related pain.⁹

International guidelines were followed to translate the original English-language version of the iHOT-33 into Thai, and no cultural adaptation was required for any part of the questionnaire. The target population in this study had selection criteria comparable to those of the target population used to develop and test the original version of the iHOT-33.

Construct validity of the TH-iHOT-33 was determined by comparing it with the HOOS and the EQ-5D-5L. The HOOS was developed to determine outcomes in patients who had hip disability with or without osteoarthritis.²⁰ The HOOS also demonstrated good psychometric properties that can be used to evaluate patients undergoing hip arthroscopy.¹¹ The HOOS was used to assess the construct validity of the Dutch version of the iHOT-33.³² The EQ-5D-5L is a generic self-report questionnaire that consists of 5 questions and 5 levels of answers. The EQ-5D

questionnaire was used to validate the German, Chinese, and Dutch versions of the iHOT-33.^{1,12,32} In this study, 6 (86%) of the 7 predefined hypotheses were confirmed. This is higher than the minimum requirement (75%) described in the COSMIN checklist,¹⁶ which indicates that the TH-iHOT-33 has good construct validity. Our results revealed that the TH-iHOT-33 has high correlation with the HOOS total score, the HOOS symptoms subscale, the HOOS pain subscale, and the EQ-5D-5L.¹⁸

The TH-iHOT-33 showed excellent internal consistency, with a Cronbach alpha of 0.98. A very high Cronbach alpha (>0.95) suggests likely redundancy of 1 or more items and is usually found in questionnaires with a large number of items.²⁹ The Cronbach alpha values of the iHOT-33 in different languages were also high, with a range within 0.90 to 0.99.^{1,12,15,26,32} In addition, the TH-iHOT-33 had high test-retest reliability, with an ICC value of 0.83. This value was higher than that of the original version¹⁵ but lower than in other languages in which the iHOT-33 has been translated.^{1,12,26,32} The SEM of the TH-iHOT-33 was 5.1 points. The SEM and SDC values of the TH-iHOT-33 in the present study were comparable to those reported by Kemp et al¹¹ and Hinman et al.⁸

Our analysis revealed that the TH-iHOT-33 had no floor or ceiling effects overall, and each item of the TH-iHOT-33 had a floor or ceiling effect ranging from 0% to <15%. These findings may be the result of the questionnaire's design, which uses a slash to mark lines. The floor and ceiling effects may be different if a Likert scale was used instead. Tijssen et al³² validated the Dutch version of the iHOT-33 and found a floor effect in 1 question and a ceiling effect in 2 questions. The MIC of the TH-iHOT-33 was 10.9 points, which was calculated according to the method described by Norman et al.²¹ This method was used in other studies^{11,32} that translated the iHOT-33. The group-level SDC of the TH-iHOT-33 was 1.9 points, and this value was less than the MIC value. The individual-level SDC was 14.2 points, which was greater than the MIC. These findings suggest that the TH-iHOT-33 is useful for group comparisons but should be used with caution when monitoring the outcomes of a patient over time.²⁹

Despite the good results concerning the validity and reliability of the TH-iHOT-33 that we observed in this study, there are some study limitations to describe. First, most of the patients who participated in this study had an activity level of 4 on a modified Tegner activity scale, which is the level of recreational sports. Competitive athletes would be expected to have different demands and perspectives relative to their hip problems, and this factor could influence floor and ceiling effects in the TH-iHOT-33 questionnaire. In addition, the calculated MIC values in this study were based on a distribution-based approach.²¹ However, there is no recommendation in the COSMIN checklist regarding the estimation of the MIC.¹⁶ Longitudinal studies would help evaluate the responsiveness of this questionnaire by comparing changes between preoperative and different postoperative time points in patients who undergo hip surgery.

CONCLUSION

The TH-iHOT-33 is a valid and reliable instrument for assessing the functional status and health-related QoL of young to middle-aged physically active Thai adults with hip disorders.

ACKNOWLEDGMENT

The authors gratefully acknowledge the original authors of the iHOT-33 for granting permission to translate the English version into Thai and Wacharapol Tapa and Narumol Sudjai for data collection and statistical analysis.

REFERENCES

- Baumann F, Weber J, Zeman F, et al. Validation of a German version of the International Hip Outcome Tool (G-iHOT33) according to the COSMIN checklist: how much improvement is clinically relevant? *Arch Orthop Trauma Surg.* 2016;136(1):83-91.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976).* 2000;25(24):3186-3191.
- Black N. Patient reported outcome measures could help transform healthcare. *BMJ.* 2013;346:f167.
- Christensen CP, Althausen PL, Mittleman MA, Lee JA, McCarthy JC. The nonarthritic hip score: reliable and validated. *Clin Orthop Relat Res.* 2003;406:75-83.
- Costa LO, Maher CG, Latimer J, et al. Clinimetric testing of three self-report outcome measures for low back pain patients in Brazil: which one is the best? *Spine (Phila Pa 1976).* 2008;33(22):2459-2463.
- Griffin DR, Parsons N, Mohtadi NG, Safran MR. A short version of the International Hip Outcome Tool (iHOT-12) for use in routine clinical practice. *Arthroscopy.* 2012;28(5):611-616.
- Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res.* 2011;20(10):1727-1736.
- Hinman RS, Dobson F, Takla A, O'Donnell J, Bennell KL. Which is the most useful patient-reported outcome in femoroacetabular impingement? Test-retest reliability of six questionnaires. *Br J Sports Med.* 2014;48(6):458-463.
- Impellizzeri FM, Jones DM, Griffin D, et al. Patient-reported outcome measures for hip-related pain: a review of the available evidence and a consensus statement from the International Hip-Related Pain Research Network, Zurich 2018. *Br J Sports Med.* 2020;54(14):848-857.
- Kamper SJ, Ostelo RW, Knol DL, et al. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. *J Clin Epidemiol.* 2010;63(7):760-766.
- Kemp JL, Collins NJ, Roos EM, Crossley KM. Psychometric properties of patient-reported outcome measures for hip arthroscopic surgery. *Am J Sports Med.* 2013;41(9):2065-2073.
- Li DH, Wang W, Li X, et al. Development of a valid Simplified Chinese version of the International Hip Outcome Tool (SC-iHOT-33) in young patients having total hip arthroplasty. *Osteoarthritis Cartilage.* 2017;25(1):94-98.
- Martin RL, Kelly BT, Philippon MJ. Evidence of validity for the Hip Outcome Score. *Arthroscopy.* 2006;22(12):1304-1311.
- Mercieca-Bebber R, King MT, Calvert MJ, Stockler MR, Friedlander M. The importance of patient-reported outcomes in clinical trials and strategies for future optimization. *Patient Relat Outcome Meas.* 2018;9:353-367.
- Mohtadi NG, Griffin DR, Pedersen ME, et al. The development and validation of a self-administered quality-of-life outcome measure for young, active patients with symptomatic hip disease: the International Hip Outcome Tool (iHOT-33). *Arthroscopy.* 2012;28(5):595-605.
- Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol.* 2010;10:22.
- Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010;63(7):737-745.
- Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J.* 2012;24(3):69-71.
- Niilsson 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 (Hoboken).* 2011;63(suppl_11):S200-S207.
- Niilsson AK, Lohmander LS, Klässbo M, Roos EM. Hip disability and Osteoarthritis Outcome Score (HOOS)—validity and responsiveness in total hip replacement. *BMC Musculoskelet Disord.* 2003;4:10.
- Norman GR, Sridhar FG, Guyatt GH, Walter SD. Relation of distribution- and anchor-based approaches in interpretation of changes in health-related quality of life. *Med Care.* 2001;39(10):1039-1047.
- Pattanaphesaj J, Thavorncharoensap M, Ramos-Gofi JM, et al. The EQ-5D-5L valuation study in Thailand. *Expert Rev Pharmacoecon Outcomes Res.* 2018;18(5):551-558.
- Polesello GC, Godoy GF, Trindade CA, et al. Translation and cross-cultural adaptation of the International Hip Outcome Tool (iHOT) into Portuguese. *Acta Ortop Bras.* 2012;20(2):88-92.
- Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med.* 2001;33(5):337-343.
- Ramisetty N, Kwon Y, Mohtadi N. Patient-reported outcome measures for hip preservation surgery—a systematic review of the literature. *J Hip Preserv Surg.* 2015;2(1):15-27.
- Ruiz-Ibán MA, Seijas R, Sallent A, et al. The international Hip Outcome Tool-33 (iHOT-33): multicenter validation and translation to Spanish. *Health Qual Life Outcomes.* 2015;13:62.
- Skendzel JG, Weber AE, Ross JR, et al. The approach to the evaluation and surgical treatment of mechanical hip pain in the young patient: AAOS exhibit selection. *J Bone Joint Surg Am.* 2013;95(18):e133.
- Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. *Clin Orthop Relat Res.* 1985;198:43-49.
- Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42.
- Thorborg K, Hölmich P, Christensen R, Petersen J, Roos EM. The Copenhagen Hip and Groin Outcome Score (HAGOS): development and validation according to the COSMIN checklist. *Br J Sports Med.* 2011;45(6):478-491.
- Thorborg K, Tijssen M, Habets B, et al. Patient-Reported Outcome (PRO) questionnaires for young to middle-aged adults with hip and groin disability: a systematic review of the clinimetric evidence. *Br J Sports Med.* 2015;49(12):812.
- Tijssen M, Tak I, Stubbe J, et al. Translation, cross-cultural adaptation, and validation of the Dutch International Hip Outcome Tool-33 (iHOT-33 NL) in young, physically active individuals with symptomatic hip joint pathology. *J Orthop Sports Phys Ther.* 2018;48(4):289-298.
- Trathitiphan W, Paholpak P, Sirichativapee W, et al. Cross-cultural adaptation and validation of the reliability of the Thai version of the Hip disability and Osteoarthritis Outcome Score (HOOS). *Rheumatol Int.* 2016;36(10):1455-1458.
- Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res.* 2005;19(1):231-240.