Relationship between iHOT12 and HOS scores in hip pain patients

Jeff Brand¹, Rich Hardy²*, Aerika Tori², Hannah Fuchs², Engin Sungur² and Emily Monroe¹

¹Sports Medicine Department, Heartland Orthopedic Specialists, 111 17th Ave E #101, Alexandria, MN 56308, USA and ²Department of Biology, Division of Science and Math, University of Minnesota Morris, Morris, MN, USA.

*Correspondence to: R. Hardy. E-mail: hardyr@morris.umn.edu

Submitted 7 May 2019; Revised 28 December 2019; revised version accepted 9 January 2020

ABSTRACT

To determine if scores of the International Hip Outcome Tool-12 (iHOT12) and the Hip Outcome Score (HOS) correlate with one another in hip pain patients. Patients reporting to an orthopedic clinic for their scheduled appointment for hip pain were given a paper survey consisting of the iHOT12 and the HOS. Demographic information [age, weight, height and body mass index (BMI)] was obtained by chart review. Overall, 114 patients were invited to voluntarily complete the surveys of which 23 declined. Our sample consisted of 91 (57 female and 34 male) patients (80% response rate). The HOS (iHOT12) explained 62% of the variation in iHOT12 (HOS) by using a linear model (Pearson’s correlation ($r$) is 0.79, $P < 0.001$). Age, weight, BMI, gender and arthritis did not show a statistically significant predictive power explaining HOS. However, only gender had a statistically significant predictive power explaining iHOT12 ($P = 0.007$). The relationship between the two scores are stronger for males ($r = 0.81, P < 0.001$) compared with females ($r = 0.77, P < 0.001$). The proportion of variations explained on one of the scores by the other are 0.66 for males and 0.59 for females. HOS score together with gender explained 64% of the variation in iHOT12 by using a linear model. iHOT12 together with the non-statistically significant gender term explained 62% of the variation in HOS by using a linear model. It may not be necessary to collect both the iHOT12 and HOS, since the predictive power of one on the other is high. Collecting HOS together with information on gender is preferable compared with collecting iHOT12. Level of evidence: Level III.

INTRODUCTION

Patient-reported outcome measures (PROMs) were originally used to assess treatment effectiveness within clinical trials [1]. However, the use of PROMs has evolved into tools that allow healthcare providers to evaluate the effects of their interventions by gaining their patients perspective in a reliable, valid and acceptable way [2]. The Hip Outcome Score (HOS) and the International Hip Outcome Tool-12 (iHOT12) are two commonly used measures for hip disabilities [3]. Both validated measures, the IHOT12 was designed to assess non-arthritic hip problems in young, active patients, while the HOS was designed to assess the treatment outcomes of arthroscopic hip surgery [4–7].

PROMS have become acceptable and routine within the healthcare industry because they are informative and observer bias is limited since patients give their perspective on their health condition [8–11]. Both healthcare providers and researchers utilize these measures to evaluate hip pathologies and the acquired data can be used for scientific investigations [12–14]. In a systematic review by Thorborg et al. [3], nine PROMs were identified for assessing young to middle-aged adults with hip disability. Although several PROMS are available, it would not be feasible or necessary to administer multiple surveys to patients due to administrative and respondent burden [9, 15].

To address this, we compared scores of two hip PROMs, the HOS and the iHOT12. Our purpose was to determine if scores of the iHOT12 and the HOS correlate with one another in hip pain patients. We hypothesized that there would be a correlation between the scores of the HOS and iHOT12.
METHODS
The University of Minnesota Institutional Review Board approved this investigation. The sample consisted of patients of the senior author, a sports fellowship trained orthopedic surgeon with 20 years of hip arthroscopy experience, from July 2017 to January 2019. Patients were included in this study if they reported for their scheduled orthopedic appointment for hip pain and were patients of the senior author and voluntarily agreed to participate in the study. Patients without hip pain, patients of another provider and patients that refused study participation were excluded. Patients that met the inclusion criteria received a packet containing an informed consent form and a survey that consisted of the HOS with ADL subscale and iHOT12 when they arrived for their appointment. After the patient voluntarily completed the informed consent form and survey by hand, demographic information [age, weight, height and body mass index (BMI)] was collected through chart review. The PROMS were then hand scored and entered into an excel spreadsheet.

A total of 114 patients were invited to voluntarily complete the surveys in which 23 refused; therefore, our final sample consisted of 91 (57 female and 34 male) patients (80% response rate). Of this sample, 25 of the 91 patients were diagnosed with osteoarthritis.

Statistical analysis
An a priori power analysis was performed using the program G*Power [16]. This analysis indicated that a sample size of 84 patients was needed to provide 80% statistical power to determine a medium effect size ($\alpha = 0.05$). Multiple linear regression analyses with indicator variable corresponding to gender were used to explain iHOT12 and HOS by using age, weight, BMI, gender and arthritis. To model the predictive power of iHOT12 and HOS on each other various linear models were used, some including gender. Also, correlation analyses by using Pearson’s correlation, $r$, was carried out used to understand the relationships between the scores of the iHOT12 and HOS, BMI and HOS scores; BMI and iHOT12 scores; age and HOS scores; and age and iHOT12 scores. Statistical software program R was utilized for all of the analyses.

RESULTS
See Table I for demographic data. The HOS (iHOT12) explained 62% of the variation in iHOT12 (HOS) by using a linear model (Pearson’s correlation ($r$) is 0.79, $P < 0.001$). Age, weight, BMI, gender and arthritis did not show a statistically significant predictive power explaining HOS (Table II). On the other hand, only gender had a ‘statistically’ significant predictive power explaining iHOT12 ($P = 0.007$). As it can be seen in Table III, the relationship between the two scores are stronger for males ($r = 0.81$, $P < 0.001$) compared with females ($r = 0.77$, $P < 0.001$). In other words, the proportion of variations explained on one of the scores by the other are 0.66 for males and 0.59 for females (Tables IV–VI). HOS score together with gender explained 64% of the variation in iHOT12 by using a linear model (Fig. 1 and Table VII). On the other hand, iHOT12 together with the non-statistically significant gender term explained 62% of the variation in HOS by using a linear model (Table VIII). iHOT12 and HOS scores for the arthritic and non-arthritic patients were not statistically different (Table IX.)

DISCUSSION
By comparing the measurement properties of the HOS with the iHOT12 in patients that reported to an orthopedic clinic complaining of hip pain, we found that these two measurements are closely related with each other. Age, weight and BMI did not show a statistically significant relationship with these measures. Our analysis showed that there is a statistically significant difference between genders for the HOS but not for iHOT12 (see Table IV). Our study suggests that it may not be necessary to collect both the iHOT12 and HOS in clinical setting, since the predictive power of one on the other is high. Collecting HOS together with information on gender is preferable compared with collecting iHOT12.

We found that the scores of the iHOT12 and the HOS correlated with one another; however, one must take into

| Table I. Demographic data |
|--------------------------|
|                        | Mean | Median | Range | SD |
|-------------------------|------|--------|-------|----|
| BMI                     | 27.41| 24.8   | 30.83 | 5.81|
| Age                     | 43.63| 45     | 64    | 16.79|
| iHOT12 scores           |      |        |       |     |
| Overall                 | 5.2  | 5.08   | 8.2   | 2.12|
| Female                  | 4.77 | 4.33   | 8.05  | 2.03|
| Male                    | 5.92 | 5.77   | 6.78  | 2.11|
| HOS scores              |      |        |       |     |
| Overall                 | 64.33| 65     | 87.5  | 19.46|
| Female                  | 61.9 | 63.24  | 87.5  | 20.02|
| Male                    | 68.41| 68.16  | 65.81 | 18.04|

BMI, body mass index; iHOT, international hip outcome tool; HOS, hip outcome score; SD, standard deviation.
consideration what patients and situations each instrument was intended for and the psychometric properties of each instrument [9]. According to Griffin et al. [4], the iHOT was developed to provide an evaluation tool for the management of non-arthritic hip problems in young, active patients for use in clinical practice to gain patient perspective for both initial assessment and postoperative follow-up care. This measure is a shorter version of the iHOT 33 with the same questionnaire characteristics with evidence of validity, reliability and responsiveness to change [4]. In contrast, including both subscales (activities of daily living and sports), the HOS has been deemed valid for use following hip arthroscopy and for use with patients enduring hip labral tears [5, 6]. There is evidence of reliability and responsiveness [7]. Martin et al. [5] reported that both of the HOS subscales had adequate internal consistency, were

Table II. The results of multiple regression analyses for predicting IHOT and HOS scores by using age, weight, IBM and gender

| Predictors       | iHOT         | HOS          |
|------------------|--------------|--------------|
|                  | Estimates    | CI           | P-value | Estimates    | CI           | P-value |
| Intercept        | 4.85         | 2.59 to 7.11 | <0.001  | 71.33       | 49.90 to 92.75 | <0.001  |
| Age              | 0.02         | −0.01 to 0.04| 0.290   | 0.03        | −0.24 to 0.31  | 0.800   |
| Weight           | 0.01         | −0.01 to 0.04| 0.251   | 0.01        | −0.20 to 0.22  | 0.917   |
| BMI              | −0.11        | −0.26 to 0.04| 0.154   | −0.43       | −1.86 to 0.99  | 0.545   |
| Sex: male        | 1.26         | 0.33 to 2.19 | 0.009   | 6.06        | −2.74 to 14.85 | 0.175   |
| Arthritis: yes   | −0.33        | −1.39 to 0.73| 0.536   | −2.73       | −12.73 to 7.28 | 0.589   |
| Observations     | 91           |              |         | 91          |              |         |
| $R^2$/$R^2$ adjusted | 0.102/0.049 | 0.043/−0.014 |         |             |             |         |

Note: Bold values are just to emphasize the statistically significant model terms.

Table III. The results of regression analysis for predicting IHOT score by using HOS score controlling for the gender

| Predictors | iHOT (female) | iHOT (male) |
|------------|---------------|-------------|
|            | Estimates     | CI          | P-value | Estimates     | CI          | P-value |
| Intercept  | −0.08         | −1.21 to 1.05 | 0.889   | −0.58         | −2.32 to 1.15 | 0.498   |
| HOS        | 0.08          | 0.06 to 0.10  | <0.001  | 0.10          | 0.07 to 0.12  | <0.001  |
| Observations | 57           |             |         | 34            |             |         |
| $R^2$/$R^2$ adjusted | 0.597/0.590 | 0.661/0.651 |         |             |             |         |

Note: Bold values are just to emphasize the statistically significant model terms.

Fig. 1. The scatter plot of iHOT12 versus HOS controlling for the gender.
Table IV. The results of the linear models to explain the impact of gender on iHOT and HOS scores

| Predictors | Estimates | CI          | P-value |
|------------|-----------|-------------|---------|
| Intercept  | 4.77      | 4.23 to 5.31| <0.001  |
| Sex: male  | 1.15      | 0.27 to 2.04| 0.011   |

| Predictors | Estimates | CI          | P-value |
|------------|-----------|-------------|---------|
| Intercept  | 61.90     | 56.82 to 66.99| <0.001  |
| Sex: male  | 6.50      | −1.81 to 14.82| 0.124   |

Observations 91

Note: Bold values are just to emphasize the statistically significant model terms.

Table V. The results of multiple regression analysis for predicting iHOT score through gender and HOS score

| Predictors | Estimates | CI          | P-value |
|------------|-----------|-------------|---------|
| Intercept  | −0.42     | −1.34 to 0.51| 0.377   |
| Sex: male  | 0.61      | 0.05 to 1.17 | 0.034   |
| HOS        | 0.08      | 0.07 to 0.10 | <0.001  |

Observations 91

\[ R^2/R^2 \text{ adjusted} = 0.644/0.635 \]

Note: Bold values are just to emphasize the statistically significant model terms.

Table VI. The results of multiple regression analysis for predicting HOS score through gender and iHOT score

| Predictors | Estimates | CI          | P-value |
|------------|-----------|-------------|---------|
| Intercept  | 26.78     | 20.12 to 33.44| <0.001  |
| Sex: male  | −1.98     | −7.35 to 3.38 | 0.464   |
| iHOT       | 7.36      | 6.13 to 8.59 | <0.001  |

Observations 91

\[ R^2/R^2 \text{ adjusted} = 0.627/0.620 \]

Note: Bold values are just to emphasize the statistically significant model terms.

Table VII. The results of regression analysis for predicting HOS score by using iHOT score

| Predictors | Estimates | CI          | P-value |
|------------|-----------|-------------|---------|
| Intercept  | −0.35     | −1.29 to 0.60 | 0.468   |
| HOS        | 0.09      | 0.07 to 0.10 | <0.001  |

Observations 91

\[ R^2/R^2 \text{ adjusted} = 0.625/0.620 \]

Note: Bold values are just to emphasize the statistically significant model terms.

Table VIII. The results of regression analysis for predicting iHOT score by using HOS score

| Predictors | Estimates | CI          | P-value |
|------------|-----------|-------------|---------|
| Intercept  | −0.92     | −1.09 to 0.75 | 0.001   |
| HOS        | 0.06      | 0.03 to 0.12 | <0.001  |

Observations 91

\[ R^2/R^2 \text{ adjusted} = 0.625/0.620 \]

Note: Bold values are just to emphasize the statistically significant model terms.

and administering both simultaneously may not provide additional information. Furthermore, collecting both measures presents a burden to the patient due to the increased number of requested answers and the time required to respond to the questions [9]. For routine clinical practice including initial assessment and postoperative follow-up, we suggest that the HOS is appropriate in this setting.

Within the literature, correlations between scores of physician-assessed and patient-assessed outcomes have been identified. Kalairajah et al. [17] compared the scores of Harris Hip Score (a physician-assessed measure) to the scores of the
Table IX. The results of the linear models to explain the impact of arthritis on iHOT and HOS scores

| Predictors            | iHOT Estimates | CI         | P-value |
|-----------------------|----------------|------------|---------|
| Intercept             | 5.28           | 4.76 to 5.80 | <0.001  |
| Arthritis: yes        | -0.28          | -1.27 to 0.72 | 0.584   |
| Observations          | 91             |            | 91      |

| Predictors            | HOS Estimates | CI         | P-value |
|-----------------------|---------------|------------|---------|
|                        | 65.44         | 60.67 to 70.20 | <0.001  |
| Arthritis: yes        | -4.01         | -13.10 to 5.09 | 0.384   |
| Observations          | 91             |            | 91      |

Note: Bold values are just to emphasize the statistically significant model terms.

Oxford hip score (a PROM) and found a good negative correlation between scores. The results of their study and of ours suggest that it may be possible to identify a single outcome measure for with patients with hip pathologies.

**Limitations**

There are several limitations to this study. First, our study surveyed patients that were experiencing hip pain in an orthopedic clinic. Our findings may be different in patients with specific injuries. Secondly, since there were more females than males, we were unable to compare data between genders. Third, participants of this study were from the Midwestern United States, which may reduce generalizability when compared with other settings. Fourth, hand-scoring of surveys may introduce measurement bias.

**CONCLUSION**

It may not be necessary to collect both the iHOT12 and HOS, since the predictive power of one on the other is high. Collecting HOS together with information on gender is preferable compared with collecting iHOT12.

**FUNDING**

Division of Science and Math, University of Minnesota Morris, Morris MN, USA.

**CONFLICT OF INTEREST STATEMENT**

None declared.

**REFERENCES**

1. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assessment* 1998; 2: i.
2. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. *J Eval Clin Pract* 2006; 12: 559–68.
3. 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: 812.
4. Griffin DR, Parsons N, Mohtadi NG et al. A short version of the International Hip Outcome Tool (iHOT-12) for use in routine clinical practice. *Arthroscopy* 2012; 28: 611–18.
5. Martin RL, Kelly BT, Philippon MJ. Evidence of validity for the hip outcome score. *Arthroscopy* 2006; 22: 1304–11.
6. Martin RL, Philippon MJ. Evidence of validity for the hip outcome score in hip arthroscopy. *Arthroscopy* 2007; 23: 822–26.
7. Martin RL, Philippon MJ. Evidence of reliability and responsiveness for the hip outcome score. *Arthroscopy* 2008; 24: 676–82.
8. Patrick DL, Burke LB, Powers JH et al. Patient-reported outcomes to support medical product labeling claims: FDA perspective. *Value Health* 2007; 10: S125–37.
9. Collins NJ, Roos EM. Patient-reported outcomes for total hip and knee arthroplasty: commonly used instruments and attributes of a “good” measure. *Clin Geriatr Med* 2012; 28: 367–94.
10. Greenhalgh J, Meadows K. The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. *J Eval Clin Pract* 1999; 5: 401–16.
11. Skevington SM, Day R, Chisholm A et al. How much do doctors use quality of life information in primary care? Testing the trans-theoretical model of behaviour change. *Qual Life Res* 2005; 14: 911–22.
12. Ohlin A, Sansone M, Ayeni OR et al. Predictors of outcome at 2-year follow-up after arthroscopic treatment of femoro-acetabular impingement. *J Hip Preserv Surg* 2017; 4: 224–230.
13. Domb BG, Dunne KF, Martin TJ et al. Patient reported outcomes for patients who returned to sport compared with those who did not after hip arthroscopy: minimum 2-year follow-up. *J Hip Preserv Surg* 2016; 3: 124–131.
14. Sansone M, Ahldén M, Jonasson P et al. Outcome of hip arthroscopy in patients with mild to moderate osteoarthritis—a prospective study. *J Hip Preserv Surg* 2016; 3: 61–7.
15. Ostendorf M, Van Stel H, Buskens E et al. Patient-reported outcome in total hip replacement: a comparison of five instruments of health status. *J Bone Joint Surg Br* 2004; 86: 801–8.
16. Faul F, Erdfelder E, Buchner A et al. Statistical power analyses using G* Power 3.1: tests for correlation and regression analyses. *Behav Res Methods* 2009; 41: 1149–60.
17. Kalajaran Y, Azurza K, Hulme C et al. Health outcome measures in the evaluation of total hip arthroplasties—a comparison between the Harris hip score and the Oxford hip score. *J Arthroplasty* 2005; 20: 1037–41.