

## RESEARCH ARTICLE

# Evaluation of the Validity and Reliability of a Persian Version of the Forgotten Joint Score in Patients Undergoing Total Hip Arthroplasty

Keyvan Ramezani, MD; Mohammad Mahdi Sarzaeem, MD; Davud Feizi, MD; Moein Bonakdar, MD; Mitra Ramezani, MD; Mahdi Sarzaeem, MD; Mohammad Movahedinia, MD

Research performed at Imam Hossein hospital, Shahid Beheshti University of Sciences, Tehran, Iran

Received: 8 May 2022

Accepted: 26 August 2023

## Abstract

**Objectives:** The increasing number of total hip arthroplasties (THA) has led to increased patient demands and expectations, making it crucial to assess patients' ability to "forget" their implants in daily life. This study aimed to determine the reliability and validity of a Persian version of the Forgotten Joint Score (P-FJS) in THA patients.

**Methods:** The questionnaire was translated bidirectionally with the permission of the questionnaire designer. Data were collected from 2018 to 2020 and included 142 patients who had undergone THA by the same surgeon at least one year ago. Participants completed the FJS questionnaire twice within a one-week interval, and the validity, reliability, and feasibility of the questionnaires were assessed using statistical tests on the HHS and OHS forms completed by all participants.

**Results:** In 142 patients (52.1% male) with a mean age of  $65 \pm 0.5$  years who answered the questionnaires, P-FJS correlated strongly with OHS and HHS. The internal consistency ( $\alpha = 0.91$ ) and reproducibility of the questionnaire were excellent. None of the floor and ceiling effects were detected.

**Conclusion:** The P-FJS questionnaire in the THA is considered a legitimate, repeatable, and self-administered survey that can be compared to its English-language counterpart. In addition, it is noteworthy that this version does not show any floor or ceiling effects.

**Level of evidence:** III

**Keywords:** Forgotten joint score, Harris hip score (HHS), OHS, P-FJS, Total hip arthroplasty (THA)

## Introduction

As the average age of the Iranian population increases, the number of total hip arthroplasty (THA) is also on the rise.<sup>1,2</sup> In recent decades, total hip arthroplasty has proven to be one of the most effective treatments for improving hip joint complications and injuries. However, evaluating the efficacy of this treatment in a patient's daily life is crucial for increasing the therapeutic efficiency of arthroplasty surgery.<sup>3,4</sup> There are numerous surgeon-dependent scoring systems and several patient-reported questionnaires to assess therapeutic efficacy, focusing mainly on pain, difficulty and quality of life.<sup>5</sup> Studies have shown that patients' demands and

expectations for recovery after arthroplasty have increased,<sup>6,7</sup> while in some cases patients' expectations are not properly recognized.

The Forgotten Joint Score (FJS) is a useful measure for evaluating the quality of recovery after THA and takes patient expectations into account without the ceiling effect.<sup>8-11</sup> The FJS focuses on the ability to forget the joint prosthesis, the main purpose of arthroplasty, which leads to the greatest patient satisfaction.<sup>12</sup> The Common Forgotten Joint Score consists of 12 questions, and each question measures the perception of the artificial joint during various daily activities.<sup>13-16</sup>

**Corresponding Author:** Keyvan Ramezani, Department of Orthopedics, Imam Hossein Medical Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Email: keyvanramezani@gmail.com



THE ONLINE VERSION OF THIS ARTICLE  
ABJS.MUMS.AC.IR

Given the importance of the FJS in recognizing the problems of patients with THA and the recent increase in THA in Iran, this study aims to evaluate the reliability and accuracy of the Persian rendition of the FJS (P-FJS) among Iranian patients undergoing total hip arthroplasty (THA).

### Materials and Methods

This observational study was conducted in the orthopedic department of a general hospital over a period from 2018 to 2020. The inclusion criteria for this study included eligible patients (n=142) who had undergone THA at least one year previously. Sample size determination was based on the recommended guideline of at least 10 patients per questionnaire item.<sup>17</sup> Patients who had undergone revision surgery or were unable to read and understand the Persian language were excluded from the study. The study was conducted in accordance with the Declaration of Helsinki, although Institutional Review Board (IRB) approval was not obtained. Informed and written consent was obtained from all participating patients. Cross-cultural adaptation of self-reported questionnaires was performed according to international guidelines.<sup>18</sup> The mean follow-up period was  $1.2 \pm 0.2$  years. Patient assessments were conducted at routine follow-up visits (n=127) by telephone (n=12), or through other social media applications (n=3).

The FJS questionnaire comprises 12 questions that assess people's awareness of their joint prosthesis during various daily activities. For each question, there are five possible answers ranging from 1 (never) to 5 (most of the time). Consequently, the raw score ranges from 12 to 60. This raw score is then converted into a scale from 0-100 and vice versa to obtain the final score. The total score is then subtracted from 100 to change the direction of the score, thus ensuring that a score of 100 represents the highest achievable score and indicates the lowest level of awareness of the joint prosthesis. If more than four answers are missing, the total score should not be used. For this reason, no patients were excluded from the study.

The translation of the questionnaire was carried out with the help of native-speaking physicians and with the permission of the questionnaire designer in a forward-backward procedure. The Persian version of the FJS (P-FJS) was subjected to a methodological review to assess validity, reliability and responsiveness according to the checklist of the consensus -based standard for the selection of health status measurement instruments (COSMIN).<sup>10</sup> The modified HHS<sup>19</sup> and the OHS<sup>20</sup> served as reference scales. Both the HHS and OHS questionnaires were previously validated for Persian speakers.<sup>21</sup> The maximum achievable score for the modified HHS was 91 points (which is an excellent result); pain was scored on a scale of 44 points (no pain) and function on a scale of 47 points (indicating excellent function). The OHS scores ranged from 12 (worst result) to 60 (best result). One week after receiving the completed questionnaires, the OHS was performed again to assess reproducibility.

### Statistical analysis

The original sample consisted of 20 subjects, including 9 men and 11 women, with an average age of  $61.4 \pm 4.7$  years (ranging from 40 to 80 years). The descriptive analyses were presented as mean, standard deviation and

percentage. To assess construct validity, Pearson r was used to examine the relationship between the HHS, P-FJS, and OHS, scoring "strong" ( $r > 0.5$ ), "moderate" ( $0.5 < r < 0.3$ ) or "weak" ( $0.3 < r < 0.1$ ) correlations.<sup>15</sup> Internal consistency was measured using Cronbach's alpha, with a correlation of  $\alpha \geq 0.90$  being considered "excellent" in terms of item consistency.<sup>13</sup> This measure indicates the longitudinal validity of the construct and identifies significant changes over time. A positive result was obtained if the smallest detectable change (SDC) was smaller than the minimum important change (MIC).

Feasibility was assessed by non-response rate, administration time, and floor and ceiling effects, which corresponded to the number of patients achieving the minimum and maximum scores, respectively. The ability of the scoring system to discriminate between extreme ranges was determined by the percentage of participants achieving the lowest or highest possible score.<sup>9</sup> In particular, the percentage of respondents who achieved the minimum or maximum score was determined. Ideally, values of more than 10% at either end of the scale should be avoided.<sup>6</sup>

Test-retest reliability or reproducibility was assessed using the intraclass correlation coefficient (ICC) and categorised as "excellent" ( $\rho > 0.75$ ), "good" ( $0.75 < \rho < 0.40$ ), or "poor" ( $\rho < 0.40$ ). The significance threshold was set at 0.05.<sup>10,17</sup>

### Results

From 2018 to 2020, a total of 142 patients underwent unilateral total hip arthroplasty (THA) for degenerative joint disorder (DJD) with subsequent osteonecrosis of the femoral head under the care of a single surgeon. Patients were closely monitored and there were no losses to follow-up, allowing data from all 142 patients to be analysed (100% response rate). The study participants had a mean follow-up time of  $1.2 \pm 0.2$  years and a mean age of  $63.0 \pm 5.8$  years. Of the total number of patients, 76 (53.5%) were male and the rest were female. The demographic data of these patients are listed in [Table 1]. The mean Patient-Reported Forgotten Joint Score (P-FJS) was  $58.18 \pm 12.41\%$  (range, 0–100%). There was no significant difference in the mean score between men and women ( $p = 0.08$ ).

Parameters	Mean/N/%	Score FJS-12	p-value
Age (year)	63.0 $\pm$ 5.8		
follow-up	1.2 $\pm$ 0.3		
male	76 (53.5%)	56.84 $\pm$ 34	0.08
female	66 (46.5%)	59.44 $\pm$ 34	
All	142(100%)	58.18 $\pm$ 12.41	

The construct validity of the P-FJS proved to be satisfactory, as it showed a strong correlation with the Oxford Hip Score (OHS) and the Harris Hip Score (HHS) [Table 2], [Figure 1]. The internal consistency of the P-FJS was excellent, with a Cronbach's alpha value of 0.93 (95% CI 0.90, 0.97). The test-retest reliability of the

questionnaire was also excellent, with an intraclass correlation coefficient (ICC) of 0.95 (95% CI 0.90, 0.99). The standard error of measurement (SEM) was determined to be 5.94, and the smallest detectable change (SDC) was 15.9. As there is no consensus on the minimum significant change (MIC) for the FJS12, this was calculated using the standard deviation divided by two, as described by Norman et al.<sup>19</sup> The MIC value was found to be 8.71, which means that the SDC value was higher than the MIC value. No floor or ceiling effect was observed. The second administration of the questionnaire was performed on average  $7.8 \pm 1.3$  days (range 7–10 days) after the first

administration to assess the test-retest reliability. The mean of the two scores was used as the final score for this study. The distribution of scores by gender is summarized in the table below [Figure 2]. The mean P-FJS score at the first administration was  $58.18 \pm 12.41\%$  and  $61.2 \pm 13.7\%$  on the second administration. The correlation between the two scores was classified as "strong" with a Pearson correlation coefficient (r) of 0.79 and a p-value of less than 0.001. The intraclass correlation coefficient was classified as "excellent" with a value of 0.79 (95% CI 0.71–0.84) and a p-value of less than 0.001.

Table 2. Mean scores and construct validity of the Persian FJS-12

Persian FJS-12 (/100)	HHS total (/91)	HHS pain (/44)	HHS function (/47)	OHS-12 (/60)
<b>58.18 ± 12.41</b>	70.6 ± 18 (30–91)	34.5 ± 6 (10–44)	36.1 ± 4 (16–47)	25.7 ± 6 (12–47)
	r = 0.62(Strong)	r = 0.59(Strong)	r = 0.63(Strong)	r = 0.51(Strong)
	p < 0.001	p < 0.001	p < 0.001	p < 0.001

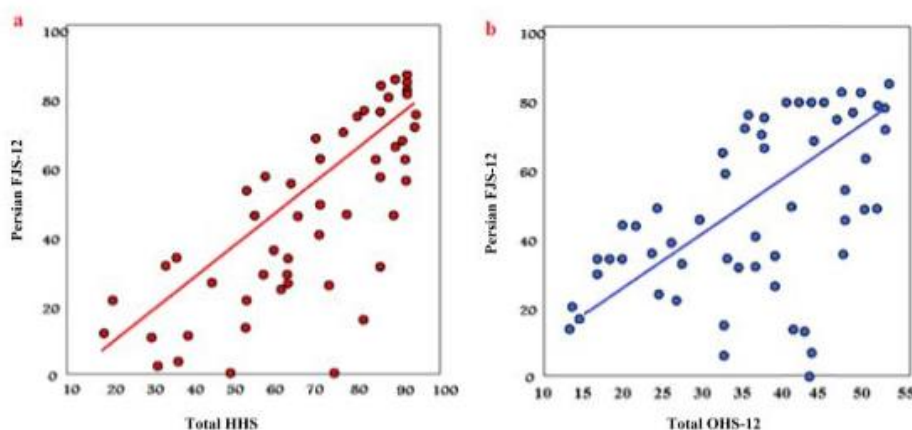


Figure 1. Regression curves between Persian FJS and total HHS (a) and OHS (b)

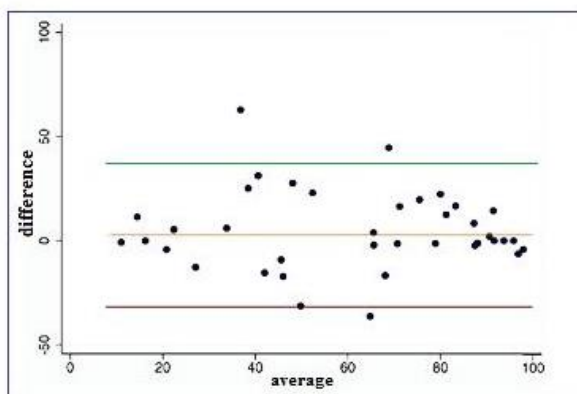


Figure 2. Reproducibility of Persian FJS: Bland-Altman plot

**Discussion**

In this study, the COSMIN method was used to examine the Persian version of the FJS, which is equivalent to the step-by-

step English version. Our results show that the P-FJS is valid, coherent and comparable to the English version in HTEP patients. The mean follow-up time was 1.2 years, with a mean FJS score of 58.18, which is consistent with other

studies.<sup>12,22</sup> In our study, the P-FJS was found to have adequate validity and reliability, with no ceiling or floor effect. In addition, our study showed good internal consistency, with Cronbach's alpha consistent with the existing literature.<sup>23,24</sup> A systematic review found that attention to FJS was significant in patients with convincing evidence of good construct validity and test-retest reliability. The study also showed moderate evidence of good internal consistency with extremely low ceiling and floor effects. Therefore, we strongly recommend the use of the FJS for the long-term assessment of the treatment of patients who are expected to reach a high level of function after THA.<sup>25</sup> In a study conducted by Larsson et al., the FJS questionnaire showed high internal consistency and reproducibility. The study also showed a low ceiling effect and positioned the FJS as a well-adapted tool for identifying patients with favorable outcomes after hip surgery.<sup>26</sup> In their study, Robinson et al. reported an average FJS score of 50.2 over a follow-up period of one year, with a negligible floor effect.<sup>27</sup> The reliability of repeat examinations has also been observed in other studies.<sup>11,22</sup>

After the first description by Behrend et al.,<sup>12</sup> the FJS was translated and validated in various languages.<sup>6,7,23,28</sup> Sethy et al. conducted a similar study in India and reported excellent compatibility of the I-FJS, as evidenced by a high Cronbach's alpha and ICC. They observed no floor or ceiling effect with the I-FJS.<sup>24</sup> Cao et al. also reported excellent Cronbach's alpha and ICC values in their study of the Chinese version of the FJS.<sup>6</sup> Thompson et al. evaluated the Danish version of the FJS questionnaire and found high internal consistency and test-retest reliability. In addition, the FJS showed good construct validity and internal consistency.<sup>29</sup> Our study showed a strong relationship between P-FJS and HHS and OHS. Similar studies have examined the relationship between the FJS and HHS and OHS in THA patients and demonstrated a direct relationship between this instrument and HHS.<sup>24</sup>

There are insufficient data on the responsiveness of FJS, and other authors have also indicated that the minimum important change (MIC) was smaller than the smallest detectable change (SDC).<sup>2,24</sup> However, these MIC values are speculative as there is no gold standard for their calculation. According to Norman et al.<sup>30</sup> the MIC for FJS was set at 10.9. In most recent studies, MIC values were around.<sup>8,9,16,22,23,28</sup>

One of the limitations of this study was the limited duration of monitoring. Hamilton et al. demonstrated that FJS is very sensitive to changes that occur between six months and one

year. However, the magnitude of the effects is much higher and there is a likelihood of variability in the results during the longer follow-up period.<sup>31</sup> A longitudinal study needs to be conducted to determine the actual changes over time. It is hoped that the survey will be translated into many languages, thereby enabling international comparison of outcomes. Despite these limitations, this study involved a considerable sample size. In addition, the comprehensive assessments were performed by a single surgeon, minimizing the potential for interobserver error.

### Conclusion

The present study showed that the Persian version of FJS in THA is a valid questionnaire and comparable to the English version. The P-FJS showed high validity, reliability, and reproducibility for hip function after THA. P-FJS is an accurate measure to identify patients with good to excellent outcomes after hip surgery, and there are no lower or upper limit effects.

### Acknowledgement

Not applicable

**Conflict of interest:** None

**Funding:** None

Keyvan Ramezani MD <sup>1</sup>  
 Mohammad Mahdi Sarzaeem MD <sup>1</sup>  
 Davud Feizi MD <sup>2</sup>  
 Moein Bonakdar MD <sup>1</sup>  
 Mitra Ramezani MD <sup>3</sup>  
 Mahdi Sarzaeem MD <sup>4</sup>  
 Mohammad Movahedinia MD <sup>5</sup>

1 Department of Orthopedics, Imam Hossein Medical Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

2 Department of Orthopedic, School of Medicine, Beasat hospital, AJA University of Medical Sciences, Tehran, Iran

3 Physical Medicine and Rehabilitation, Isfahan University of Medical Sciences, Isfahan, Iran

4 Medical student, Shahid Beheshti University of Medical Sciences, Tehran, Iran

5 Shahid Beheshti University of Medical Sciences, Tehran, Iran

### References

- Aslani H, Nourbakhsh ST, Lahiji FA, et al. Iranian Joint Registry (Iranian National Hip and Knee Arthroplasty Registry). *Arch Bone Jt Surg*.2016; 4(2):192-6.
- Tolppanen AM, Lavikainen P, Soinen H, Hartikainen S. Incident hip fractures among community dwelling persons with Alzheimer's disease in a Finnish nationwide register-based cohort. *PLoS One*. 2013; 8(3):e59124. doi: 10.1371/journal.pone.0059124.
- Hardt S, Hube R, Perka C. Total Hip Arthroplasty for High Hip Dislocation. *Z Orthop Unfall*.2020; 158(2):170-183. doi: 10.1055/a-0946-2750.
- Quitmann H. Supercapsular percutaneously assisted (SuperPath) approach in total hip arthroplasty: Surgical technique and preliminary results. *Oper Orthop Traumatol*.2019; 31(6):536-546. doi: 10.1007/s00064-019-0597-5.

5. Bullinger M, Alonso J, Apolone G, et al. Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. *International Quality of Life Assessment. J Clin Epidemiol.*1998; 51(11):913-23. doi: 10.1016/s0895-4356(98)00082-1.
6. Cao S, Liu N, Han W, et al. Simplified Chinese version of the Forgotten Joint Score (FJS) for patients who underwent joint arthroplasty: cross-cultural adaptation and validation. *J Orthop Surg Res.*2017; 12(1):6. doi: 10.1186/s13018-016-0508-5.
7. Carlson VR, Post ZD, Orozco FR, Davis DM, Lutz RW, Ong AC. When Does the Knee Feel Normal Again: A Cross-Sectional Study Assessing the Forgotten Joint Score in Patients After Total Knee Arthroplasty. *J Arthroplasty.*2018; 33(3):700-703. doi: 10.1016/j.arth.2017.09.063.
8. Marx RG, Jones EC, Atwan NC, Closkey RF, Salvati EA, Sculco TP. Measuring improvement following total hip and knee arthroplasty using patient-based measures of outcome. *J Bone Joint Surg Am.*2005; 87(9):1999-2005. doi: 10.2106/JBJS.D.02286.
9. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res.*1995; 4(4):293-307. doi: 10.1007/BF01593882.
10. Mokkink LB, Terwee CB, Knol DL, et al. Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. *BMC Med Res Methodol.*2006;6:2. doi: 10.1186/1471-2288-6-2.
11. Na SE, Ha CW, Lee CH. A new high-flexion knee scoring system to eliminate the ceiling effect. *Clin Orthop Relat Res.*2012; 470(2):584-93. doi: 10.1007/s11999-011-2203-5.
12. Behrend H, Giesinger K, Giesinger JM, Kuster MS. The "forgotten joint" as the ultimate goal in joint arthroplasty: validation of a new patient-reported outcome measure. *J Arthroplasty.*2012; 27(3):430-436.e1. doi: 10.1016/j.arth.2011.06.035.
13. Clement ND, Weir D, Holland JP, Gerrand CH, Deehan DJ. An Overview and Predictors of Achieving the Postoperative Ceiling Effect of the WOMAC Score Following Total Knee Arthroplasty. *J Arthroplasty.*2019; 34(2):273-280. doi: 10.1016/j.arth.2018.10.004.
14. Eichler D, Beaulieu Y, Barry J, Massé V, Vendittoli PA. Perception of a Natural Joint after Total Knee Arthroplasty. *J Arthroplasty.*2020; 35(2):358-363. doi: 10.1016/j.arth.2019.09.036.
15. Fleiss JL, Levin B, Paik MC, eds. *Statistical methods for rates and proportions.* 3rd ed. John Wiley & Sons; 2013.
16. Gandek B. Measurement properties of the Western Ontario and McMaster Universities Osteoarthritis Index: a systematic review. *Arthritis Care Res (Hoboken).*2015; 67(2):216-29. doi: 10.1002/acr.22415.
17. 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. doi: 10.1016/j.jclinepi.2006.03.012.
18. Beaton DE, Bombardier C, Guillemin F, Ferraz MB, Howie CR, Giesinger K. Responsiveness and ceiling effects of the Forgotten Joint Score-12 following total hip arthroplasty. *Bone Joint Res.*2016; 5(3):87-91. doi: 10.1302/2046-  
for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976).*2000; 25(24):3186-91. doi: 10.1097/00007632-200012150-00014.
19. Byrd JW, Jones KS. Prospective analysis of hip arthroscopy with 2-year follow-up. *Arthroscopy.*2000; 16(6):578-87. doi: 10.1053/jars.2000.7683.
20. Delaunay C, Epinette JA, Dawson J, Murray D, Jolles BM. Cross-cultural adaptations of the Oxford-12 HIP score to the French speaking population. *Orthop Traumatol Surg Res.*2009; 95(2):89-99. doi: 10.1016/j.otsr.2009.01.003.
21. Nourbakhsh M, Zarezadeh A, Shemshaki H, Etemadifar MR, Moezi M, Mazoochian F. Translation and cultural adaptation of the oxford hip score for Iranian population. *Int J Prev Med.*2013; 4(2):141-5.
22. Shadid MB, Vinken NS, Marting LN, Wolterbeek N. The Dutch version of the Forgotten Joint Score: test-retesting reliability and validation. *Acta Orthop Belg.*2016; 82(1):112-8.
23. Klouche S, Giesinger JM, Sariali EH. Translation, cross-cultural adaptation and validation of the French version of the Forgotten Joint Score in total hip arthroplasty. *Orthop Traumatol Surg Res.*2018; 104(5):657-661. doi: 10.1016/j.otsr.2018.04.010.
24. Sethy SS, Goyal T, Paul S, Das SL, Choudhury AK, Kalia RB. Translation and Validation of Forgotten Joint Score for Total Hip Arthroplasty for Indian Population. *Indian J Orthop.*2020; 54(6):892-900. doi: 10.1007/s43465-020-00228-x.
25. Adriani M, Malahias MA, Gu A, Kahlenberg CA, Ast MP, Sculco PK. Determining the Validity, Reliability, and Utility of the Forgotten Joint Score: A Systematic Review. *J Arthroplasty.*2020; 35(4):1137-1144. doi: 10.1016/j.arth.2019.10.058.
26. Larsson A, Rolfson O, Kärrholm J. Evaluation of Forgotten Joint Score in total hip arthroplasty with Oxford Hip Score as reference standard. *Acta Orthop.*2019; 90(3):253-257. doi: 10.1080/17453674.2019.1599252.
27. Robinson PG, Rankin CS, Murray IR, Maempel JF, Gaston P, Hamilton DF. The forgotten joint score-12 is a valid and responsive outcome tool for measuring success following hip arthroscopy for femoroacetabular impingement syndrome. *Knee Surg Sports Traumatol Arthrosc.*2021; 29(5):1378-1384. doi: 10.1007/s00167-020-06138-7.
28. Hamilton DF, Loth FL, Giesinger JM, et al. Validation of the English language Forgotten Joint Score-12 as an outcome measure for total hip and knee arthroplasty in a British population. *Bone Joint J.*2017; 99-B (2):218-224. doi: 10.1302/0301-620X.99B2.BJJ-2016-0606.R1.
29. Thomsen MG, Latifi R, Kallemose T, Barfod KW, Husted H, Troelsen A. Good validity and reliability of the forgotten joint score in evaluating the outcome of total knee arthroplasty. *Acta Orthop.*2016; 87(3):280-5. doi: 10.3109/17453674.2016.1156934.
30. Norman GR, Wyrwich KW, Patrick DL. The mathematical relationship among different forms of responsiveness coefficients. *Qual Life Res.*2007; 16(5):815-22. doi: 10.1007/s11136-007-9180-x.
31. Hamilton DF, Giesinger JM, MacDonald DJ, Simpson AH, 3758.53.2000480.