The Journal of Arthroplasty xxx (2019) 1–7

Contents lists available at ScienceDirect

The Journal of Arthroplasty

journal homepage: www.arthroplastyjournal.org

Can We Help Patients Forget Their Joint? Determining a Threshold for Successful Outcome for the Forgotten Joint Score

Philip J. Rosinsky, MD, Jeffrey W. Chen, BA, Ajay C. Lall, MD, MS, Jacob Shapira, MD, David R. Maldonado, MD, Benjamin G. Domb, MD *

American Hip Institute, Des Plaines, IL

ARTICLE INFO

Article history:
Received 7 June 2019
Received in revised form
8 July 2019
Accepted 4 August 2019
Available online xxx

Keywords:
total hip arthroplasty
threshold
outcome
Forgotten Joint Score
accuracy

ABSTRACT

Background: Clinically important thresholds improve interpretability of patient-reported outcomes. A threshold for a successful outcome does not exist for the Forgotten Joint Score (FJS). The purpose of this study is to determine a threshold score for the FJS, 1 and 2 years after total hip arthroplasty (THA).
Methods: A retrospective analysis of 247 primary THA recipients between December 2012 and April 2017 was performed. A binary "successful treatment" was defined as achieving a composite criterion of pain, function, and satisfaction. Receiver operator characteristic analysis determined thresholds for successful outcome at 1 and 2 years postoperatively, subanalyzed by demographics. Results were validated by a 75th centile comparison. The ceiling effect of FJS was also assessed.
Results: The average FJS was 70.06 ± 29.39 and 75.05 ± 28.73 at 1 and 2 years, respectively ($P < .001$). The proportion of patients meeting the composite criteria for success was 66.8% at 1 year and 76.5% at 2 years ($P = .017$). The receiver operator characteristic analysis for FJS at 1 and 2 years yielded excellent accuracy as defined by area under the curve (0.91 and 0.92, respectively). The threshold values were 73.96 and 69.79 at the respective time points. A mild ceiling effect was found with 16% and 23% of cases achieving a score of 100 at 1 and 2 years, respectively.
Conclusion: The FJS has excellent accuracy in demonstrating successful outcome following THA. The FJS threshold for success at 1 and 2 years postoperatively is 73.96 and 69.79, respectively. The higher rates of success at 2 years, along with a rise in the mean FJS, may indicate continued clinical improvement up to 2 years after THA.

With the shift toward patient-centered care, patient-reported outcome (PRO) scores have become a central facet of post-operative outcome assessment in patients undergoing hip arthroplasty. Several well-validated PROs, such as the 36-Item Short Form Survey, Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), and the Oxford Hip Score (OHS), have been designed to evaluate general health-related quality-of-life as well as hip-specific measures [1].

One of the primary goals of joint arthroplasty is to restore the "natural feeling" of the joint. This goal of reducing joint awareness is a more complex outcome than strict measures of pain and function. It requires a comprehensive consideration of physical and psychological perception of the joint as a foreign body [2]. The Forgotten Joint Score (FJS) has been developed as a PRO to assess joint awareness and has increasingly been used in both hip and knee arthroplasty literature. First reported by Behrend et al [3], the FJS has since been validated in multiple languages [4–10]. The FJS was originally proposed to possess a lower ceiling effect than other PROs, enabling it to better discern between patients achieving good and excellent results.

Although PRO scores can be valuable in comparison of surgical treatments and differences between different surgical population groups, clinical interpretation of these differences may be misleading. To improve interpretability of PROs and contextualize these scores, different clinically meaningful cutoff points have been established [11,12]. The most commonly reported thresholds are

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to https://doi.org/10.1016/j.arth.2019.08.014.

* Reprint requests: Benjamin G. Domb, MD, American Hip Institute, 999 E Touhy Ave, Suite 450, Des Plaines, IL 60018.

https://doi.org/10.1016/j.arth.2019.08.014
0883-5403/

the minimal clinical important difference, substantial clinical benefit, and the patient acceptable symptom state (PASS). These thresholds for various scores have increased the clinical relevance of PROs by serving as benchmarks for treatment success.

Although a successful threshold exists for other outcome scores of hip arthroplasty [11,13], one does not exist for the FJS. To improve the clinical applicability of the FJS, there is a need to establish a threshold for successful outcome. Therefore, the purpose of this study was (1) to determine a threshold score for the FJS at 1 and 2 years following total hip arthroplasty (THA) and (2) to assess the ceiling effect of FJS at these timeframes.

Patients and Methods

Sample Population

This retrospective analysis was conducted on a single-center, single surgeon's hip arthroplasty registry for consecutive patients undergoing primary THA by the senior author (BGD) between December 2012 and April 2017 who completed PRO questionnaires. Patients were included in the study if questionnaires were completed at both 1 and 2 years following THA. Patients who were of workers' compensation status were excluded. THAs were performed either by direct anterior or posterior approach. Robotic-assisted cases were performed using the MAKO Robotic-Arm (Stryker Corporation, Kalamazoo, MI) for acetabular reaming, cutting, and component placement, while nonrobotic cases were performed under fluoroscopic guidance. All THAs were performed using a cementless cup with polyethylene liner (Trinity; Corin, UK or Mpact; Medacta, Switzerland), uncemented stem (Metafix; Corin, UK or Quadra; Medacta, Switzerland), and a ceramic head (Biolox Delta; CeramTec, Plochingen, Germany).

Although the present study presents novel findings, data on some patients may have been used for other studies. All data collection received Institutional Review Board approval.

Outcome Measures

Definition of Treatment Success

In order to define a successful outcome to use as an external criterion for threshold determination, a combination of the following was used:

1. Average daily pain, as reported by patients on a 5-item Likert-graded scale (ranging from "none" to "severe"). This study considered patients reporting either "none" or "slight" levels of pain, the 2 lowest levels, as a successful outcome. The choice of the Likert-graded evaluation of pain was based on the recommendation of the Patient-Reported Outcome Measures Working Group of the International Society of Arthroplasty Registries [14].
2. Functional level was determined based on functional assessment using the gait and activities of daily living (ADL) domains of the Harris Hip Score (HHS). Patients were considered to have a successful outcome only if they reported perfect function for the ADL questions, and perfect or nearly perfect function on gait questions.
3. Satisfaction was graded on a Visual Analog Scale ranging from 0 to 10. Satisfaction $\geq$5 was defined as successful based on previous literature which used this cutoff for determination of treatment success following joint arthroplasty [15,16].

This study considered THA treatment success as achievement of all 3 components: demonstrating no or minimal pain, high functional level, and overall high satisfaction. This created a dichotomous definition of treatment success that was used as the binary external criterion for the receiver operator characteristic (ROC) analysis.

The Forgotten Joint Score

The FJS was constructed to measure the success of THA with the decided aim of assessing the ability of the patient to forget about their prosthetic implant. The FJS integrates a variety of variables such as pain, stiffness, function in ADL, patients' expectations, patients' activity levels, and psychosocial factors [3]. The FJS is scored on a scale of 0-100, with higher scores denoting lower joint awareness. The FJS score has been purported to have a lower ceiling effect when compared to other hip-specific outcome measures such as WOMAC and OHS [4,17–19].

Statistical Analysis

Demographic characteristics were reported as means, medians, and standard deviations for continuous variables, and counts and percentages for categorical variables.

Determination of thresholds was based on ROC analysis using the FJS as the predictor for the dichotomous variable "treatment success" as previously outlined. The area under the curve (AUC) accompanied by 95% confidence intervals (CIs) was used as a measure of diagnostic accuracy. An AUC = 0.50 was equated with random assignment, while AUC = 1.0 as perfectly accurate prediction. Values between 0.70 and 0.80 were considered acceptable discrimination and AUC above 0.80 was considered excellent discrimination [20]. The specificity and sensitivity were reported for each ROC curve. The Youden's index was used to indicate the point on the curve which achieves the highest sensitivity and specificity as the optimal cutoff value (threshold). This threshold as determined by the Youden index can be interpreted as the maximal effectiveness of the score's ability to differentiate between a dichotomous outcome. The validated 75th centile approach as described by Tubach et al [21] was used to cross reference the cutoff found by the Youden's index. This approach uses the 75th centile of scores of patients who achieved treatment success to approximate the threshold for success, and has been shown to be within 95% CI of the ROC model [22]. De Long's test was used for pair-wise comparisons of ROC curves and thresholds at 1 and 2 years [23,24].

Statistical analysis was performed with Python (Python Foundation Version 3.7) and R (R Statistical Software Version 3.5.2). Data were assessed for normality using the Shapiro-Wilk test and was assessed for equal variance using the *F*-test. The Mann-Whitney and Welch tests were used to compare non-normally distributed data with equal and unequal variances, respectively. Normally distributed data sets were compared using the Student's *t*-test. Categorical variables were assessed with the chi-squared test or Fisher's exact test. The threshold for statistical significance was set at $P < .05$.

Results

Patient Demographics

A total of 247 primary THA patients were included in this study. All patients had completed PRO questionnaires at both 1 and 2 years following THA. Patient demographics and percentage of

Table 1
Patient Demographics.

Demographic	
Age at surgery (y) (mean ± SD)	58.51 ± 9.77
Body mass index (kg/m^2) (mean ± SD)	29.52 ± 5.68
Gender (male:female)	(106:141)
Approach (anterior:posterior)	(219:28)
Guidance (robotic:fluoroscopic)	(223:14)

SD, standard deviation.

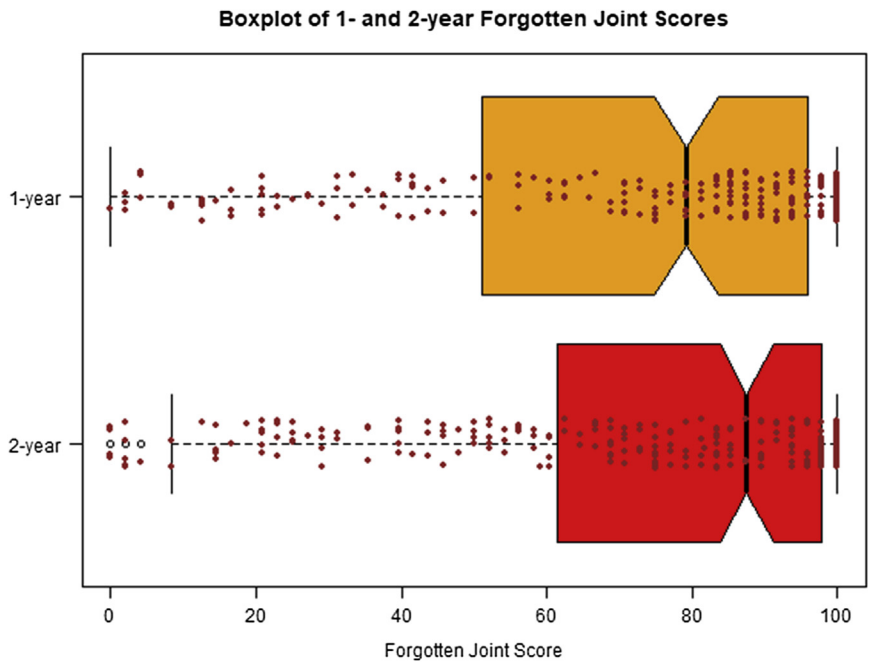

Fig. 1. Box and whisker plot of the 1-year and 2-year Forgotten Joint Scores. The box spans the first to third quartile, with the median line at the center. The whiskers span 1.5 times the interquartile range. The notches indicate the 95% confidence interval around the median. A jitter plot of each respondent is overlain on top of the box and whiskers.

patients satisfying each of the "success" criteria are presented in Table 1.

The mean FJS was 70.06 ± 29.39 (median = 79.17) and 75.05 ± 28.73 (median = 87.50) at 1 and 2 years, respectively (Fig. 1, Table 2). Although the difference between the FJS at 1 and 2 years was significantly different ($P < .001$), the effect size was small to medium (paired Cohen's $d = 0.259$, CI = 0.081-0.436) [25].

ROC Analysis and Thresholds

The proportion of patients meeting the composite criteria for satisfaction, pain, and function was 165 of 247 (66.8%) at 1 year, and 189 of 247 (76.5%) at 2 years ($P = .017$). At both 1 and 2 years, the ROC curves exhibited high AUCs (>0.8) (AUC = 0.91, CI = 0.87-0.95 and AUC = 0.92, CI = 0.87-0.96), demonstrating excellent discrimination for the FJS between patients attaining and not attaining treatment success (Fig. 1). The specificity and sensitivity of the ROC curves were high, with specificities of 0.8448 and 0.8902 and sensitivities of 0.8783 and 0.7879 at the 1-year and 2-year marks, respectively. The Youden's index found the threshold score at 1 year to be 73.96 and at 2 years to be 69.79 ($P = .471$) (Fig. 2).

Furthermore, ROCs were assessed for each demographic subgroup (gender, age, and body mass index) (Table 3). The ROC-AUC across all demographics was within excellent ranges with no significant differences between the curves or thresholds.

The 75th centile approach of the cumulative function of patients achieving treatment success found the threshold for FJS at 1 year to be 75 and the threshold at 2 years to be 81 (Fig. 3).

Assessment of Ceiling Effect

At 1 year, 40 of 247 (16.2%) patients achieved FJS of 100, while at 2 years this number increased to 57 of 247 (23.0%) ($P = .054$). Figure 4 shows the distribution of the FJS at 1-year and 2-year marks.

Complications and Revisions

At the 2-year follow up, there were 5 cases (2%) that had required a revision THA, 3 other cases (1.2%) that had required subsequent surgery due to infection or periprosthetic fracture, and 25 cases (10.1%) that suffered minor complications, most of which resolved by the second year follow-up (Table 4).

Discussion

The purpose of this study is to determine a threshold for success for the FJS based on attainment of high functional activity level, minimal pain, and high satisfaction following THA. This study found high accuracy of the FJS at predicting treatment success (ROC-AUC > 0.9) at both 1 and 2 years following THA. The FJS thresholds were 73.96 and 69.79 at 1 and 2 years, respectively. Although the values found for the thresholds seem to change between the first and second year, this difference was not statistically significant, with overlapping CIs. This study also found a time-dependent ceiling effect, with 16.2% of patients reporting maximal FJS at 1 year, and 23% reporting maximal FJS at 2 years.

Giesinger et al [26] conducted a study based on an online survey of 2017 individuals of the general US population in order to determine normative FJS-12 scores. In their study, in which only 1.4% of respondents report a history of THA, the mean FJS-12 was 70.6 (69.9 for females and 71.2 for males). In a previous study,

Table 2
Outcome Response at 1 and 2 Y Post-Operatively.

Demographic	1 Y	2 Y	*P*-Value
FJS (mean ± SD)	70.06 ± 29.39	75.05 ± 28.73	<.001
Patients meeting success criteria			
Pain (n, %)	171 (69.23)	195 (78.95)	.01
Satisfaction (n, %)	228 (92.31)	233 (94.33)	.37
Function (n, %)	214 (86.64)	216 (87.45)	.79
Total anchor (n, %)	165 (66.8)	189 (76.52)	.02

FJS, Forgotten Joint Score; SD, standard deviation.

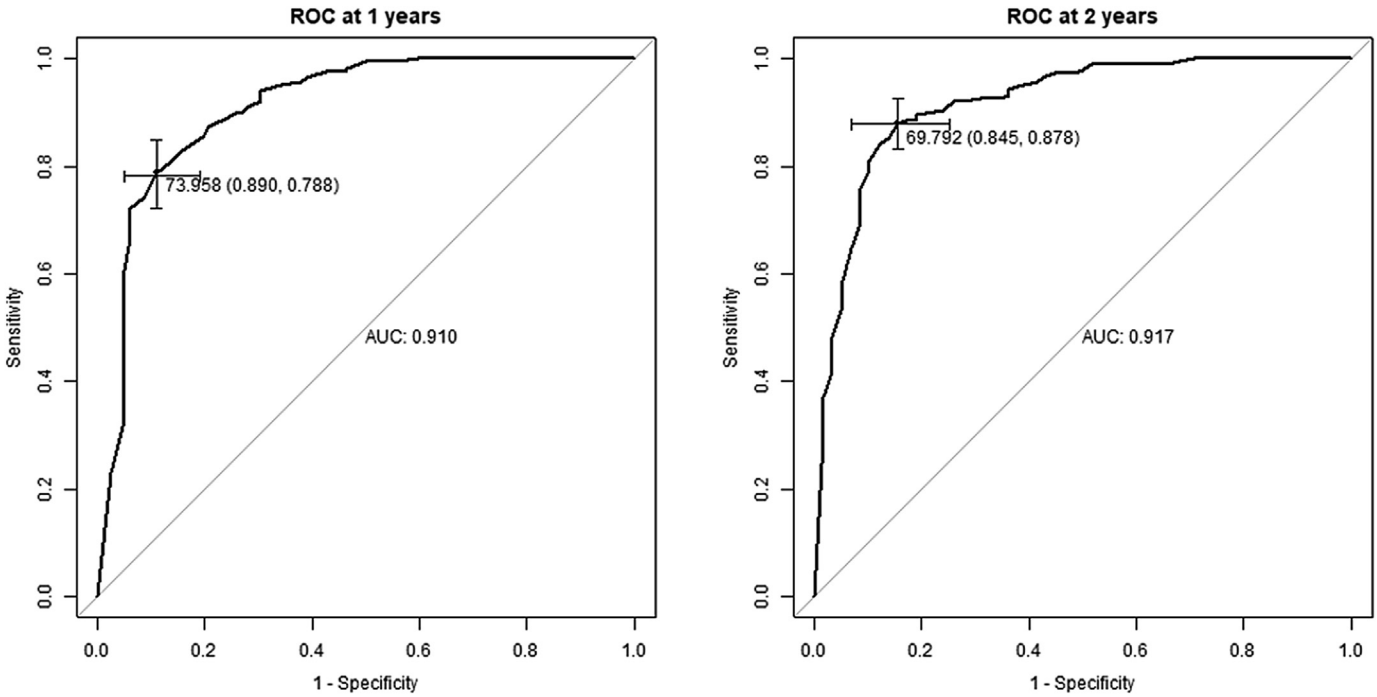

Fig. 2. Receiver operator characteristic curves at 1 and 2 years of all patients undergoing arthroplasty. The highest threshold as calculated by the Youden index is labeled with specificity and sensitivity noted in parentheses. The corresponding 95% confidence intervals are plotted around for the threshold. AUC, area under the curve; ROC, receiver operator characteristic.

Giesinger et al [20] suggested using general population norms as a way of validating "success" thresholds of scores. In the current study, the average FJS achieved at 1 year was similar to the mean reported by Giesinger et al for the non-THA adult population. The average FJS at 2 years, however, was slightly higher. This shows that the average patient following THA achieves similar joint awareness as the general population with a native joint. Additionally, the thresholds which were detected in this study (69.79 and 73.96) are comparable to the general, non-THA population, attesting to the validity of these thresholds as successful cutoff points to determine successful outcome following THA.

Puliero et al [27] studied the correlation between 3 scores: the Patient Joint Perception (PJP), the WOMAC, and the FJS. Although they did not determine a threshold based on ROC analysis, in their cohort, patients with an "artificial joint with minimal restrictions" as defined by the PJP score had an average WOMAC of 14.6 and an average FJS of 79. These scores are slightly better than previous PASS thresholds determined for WOMAC [28], and slightly higher than the FJS threshold as determined in the present study. One explanation for this is that while the present study performed a true threshold analysis based on ROC curves, the study by Puliero et al only determined the average WOMAC or FJS of patients in a specified category of PJP. Another possible reason for this difference may be that in the study by Puliero et al, patients had a mean follow-up of 68 months, but patients with less than 4-year follow-up were excluded. These excluded patients may be those with less successful outcomes, thereby skewing the data toward more superior outcomes.

Table 3
Receiver Operator Characteristic Analysis at 1 and 2 Y Post-Operatively.

Demographic	1 Y						2 Y					
	Threshold	Sensitivity	Specificity	AUC	Ceiling, n (%)	PASS, n (%)	Threshold	Sensitivity	Specificity	AUC	Ceiling, n (%)	PASS, n (%)
All (n = 247)	73.96	0.79	0.89	0.91 (0.87-0.95)	40 (16)	139 (56)	69.79	0.88	0.84	0.92 (0.87-0.96)	57 (23)	175 (71)
Male (n = 106)	69.79	0.82	0.83	0.86 (0.76-0.95)	17 (16)	67 (63)	80.21	0.76	0.89	0.91 (0.85-0.97)	28 (26)	69 (65)
Female (n = 141)	76.04	0.79	0.94	0.94 (0.9-0.98)	23 (16)	73 (52)	71.88	0.87	0.93	0.92 (0.86-0.98)	29 (21)	91 (65)
Age, tercile 1 (n = 82)	69.79	0.89	0.88	0.94 (0.88-1)	14 (17)	53 (65)	77.08	0.76	1.00	0.94 (0.89-0.99)	22 (27)	50 (61)
Age, tercile 2 (n = 81)	53.13	0.93	0.85	0.96 (0.92-1)	15 (19)	55 (68)	73.96	0.89	0.95	0.97 (0.93-1)	17 (21)	55 (68)
Age, tercile 3 (n = 84)	71.88	0.80	0.83	0.84 (0.74-0.94)	11 (13)	48 (57)	71.88	0.87	0.77	0.86 (0.76-0.96)	18 (21)	59 (70)
BMI, tercile 1 (n = 81)	70.83	0.82	0.94	0.93 (0.87-0.99)	13 (16)	43 (53)	69.79	0.81	1.00	0.95 (0.9-0.99)	18 (22)	52 (64)
BMI, tercile 2 (n = 83)	69.79	0.85	0.87	0.92 (0.84-1)	11 (13)	54 (65)	80.21	0.76	0.86	0.87 (0.79-0.96)	17 (20)	50 (60)
BMI, tercile 3 (n = 83)	73.96	0.80	0.86	0.88 (0.79-0.97)	16 (19)	48 (58)	69.79	0.95	0.90	0.94 (0.85-1)	22 (27)	62 (75)

AUC, area under the curve; PASS, patient acceptable symptomatic state; BMI, body mass index.

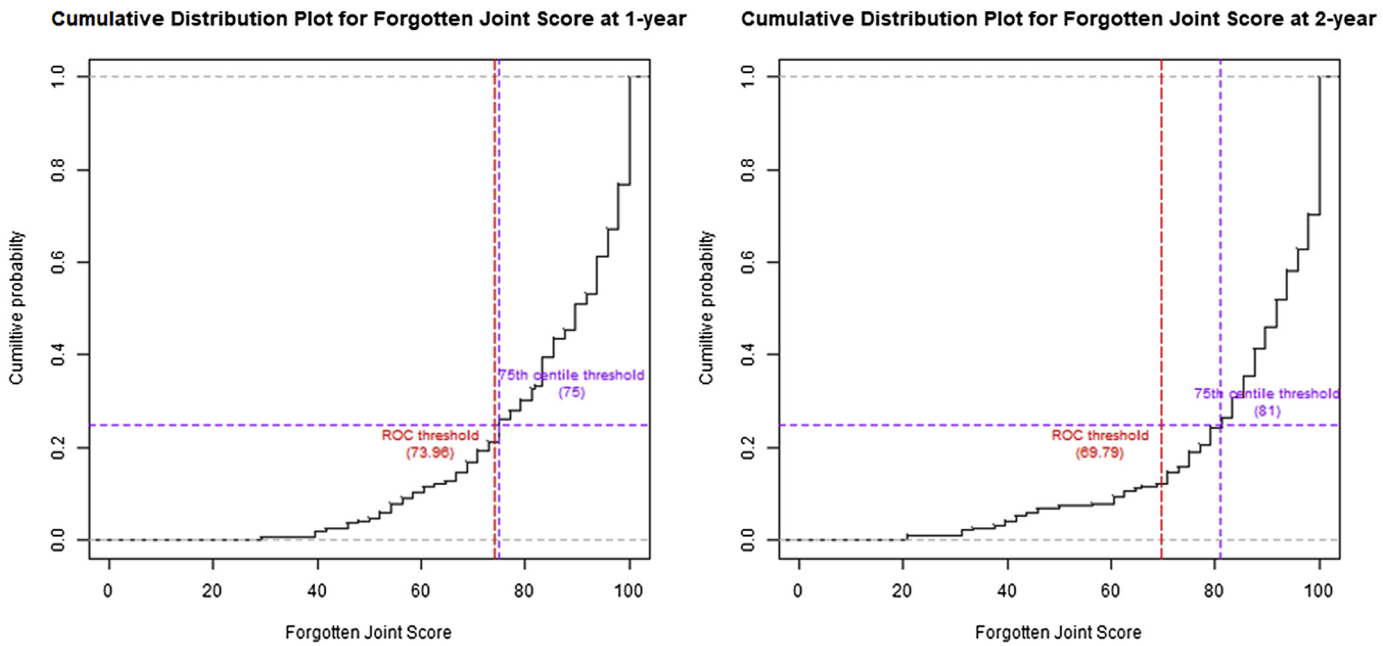

Fig. 3. The 75th centile of all responders at 1 and 2 years. The red vertical lines indicate the calculated threshold by the ROC Youden index. The purple lines show the 75th centile of responders and the corresponding Forgotten Joint Score.

Various methods exist for determining cutoff points for treatment success. The most common and accepted method is an anchor-based approach. The anchor-based approach requires a binary external criterion, constructed by either a single outcome or a composite of outcomes. The threshold identified in this study is not a true PASS score, but a more comprehensive threshold. A classic PASS is determined by analysis of the patients' answer to the question: "Do you deem your condition acceptable?" [11] Previous studies on joint arthroplasty have suggested that different anchors should be used for each of the domains of pain, function, and satisfaction [13,20,29].

The high ROC-AUC signifies that the FJS had a high accuracy in discriminating between patients who attained high satisfaction, were pain-free, and with high-functional state from those who did not reach these benchmarks. This finding was true at both the 1-year and 2-year mark, as well as across different demographic subgroups. Similar studies on score thresholds in hip and knee arthroplasty have generally reported lower ROC-AUCs. Giesinger et al [20] reported AUCs of 0.69-0.83 for various knee PROs. Keurentjes et al [16] reported on PASS thresholds for the OHS and Oxford Knee Score and reported AUCs of 0.83 and 0.72, respectively. In another study on OHS and Oxford Knee Score thresholds, Judge et al

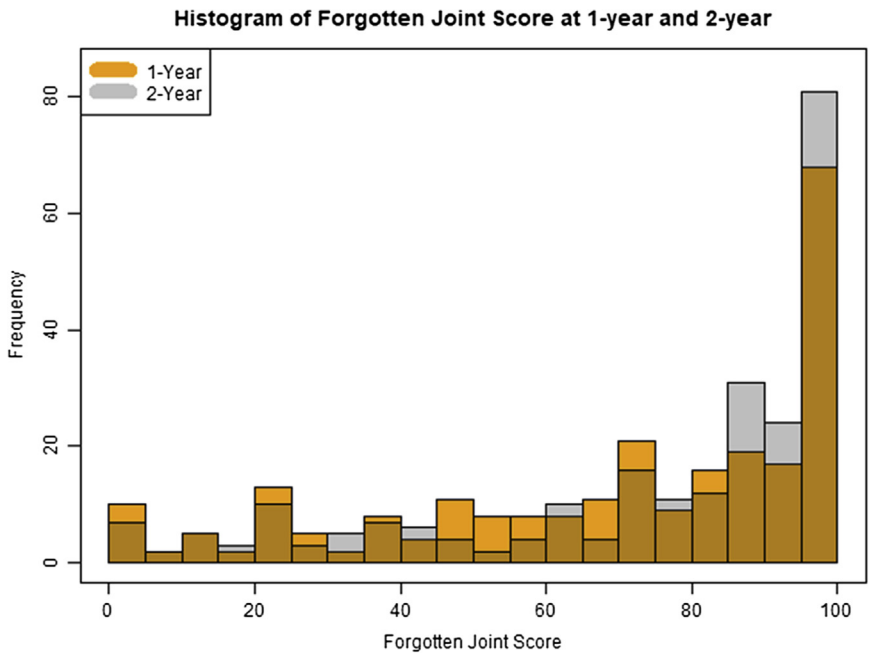

Fig. 4. Histogram of Forgotten Joint Scores at 1 and 2 years.

Table 4
Complications and Revision Surgery.

Complication		Number (%)
Revisions	Aseptic loosening	4 (1.6%)
	Periprosthetic fracture	1 (0.4%)
Wound healing delay and infections	Deep infection: irrigation and debridement	2 (0.8%)
	Superficial: nonoperative	8 (3.2%)
Periprosthetic fracture[a]		2 (0.8%)
Numbness		14 (5.6%)
Others	Iliopsoas tendinitis, central retinal vein occlusion, exacerbation of multiple sclerosis	3 (1.2%)

[a] One patient underwent revision of the femoral component and 1 patient underwent open reduction and internal fixation of the fracture.

[15] reported AUCs ranging from 0.8 to 0.86. Although the high AUC found in our study reflects the accuracy of discriminating successful outcomes, the fact that the high AUC held true at different time points and across demographic subgroups is important in assessing the content validity of the FJS.

Most studies evaluating the time-dependent trend of PROs following THA have focused on intervals up to 1 year postoperatively. Few studies have assessed changes up to 2 years after surgery. A systematic review by Browne et al [30] showed a continued improvement in PROs following THA for at least 1 year. Naal et al [31] showed that the PASS score increased over time, at least up to 1 year postoperatively. The current study found that the mean FJS improved from 1 year to 2 years after surgery (70.06-75.5, $P < .001$), with a higher proportion of patients reporting attaining successful outcome (76.5% vs 66.8%, $P = .017$). This may be related to the nature of the FJS score, which was designed to assess joint awareness. Hamilton et al [18] showed that the FJS has a larger change between 6 and 12 months compared to the OHS, which showed minimal changes. One may postulate that as time elapses after surgery, patients with successful outcomes become less aware of the presence of their artificial joints, hence the rise in the mean FJS scores between 1 and 2 years. This continued improvement in joint awareness contrasts other studies comparing PROs longitudinally, which have generally shown limited improvement beyond 1 year.

As outcomes of THA have improved over the last decades and more THA is being undertaken in younger populations, many of the PROs used to evaluate THA have exhibited ceiling effects [1,3,32]. A ceiling effect for a PRO is considered as more than 15% of a population scoring the highest possible score [33]. Behrend et al proposed the FJS as a PRO with a lower ceiling effect compared to other traditional PROs, with a number of studies backing their findings [3,18,34]. However, other studies have demonstrated less favorable results for the FJS, showing ceiling effects up to 30% [17,27]. This study did not compare FJS to other PROs, though it did find that a mild ceiling effect does exist, especially at 2 years (23%). Higher ceiling effects in hip arthroplasty may either be due to a limitation of the PRO ability to evaluate outcomes or to genuine success of THA as a procedure.

Strengths

Defining a success threshold for treatment requires setting an external anchor. Although many studies use a single criterion to determine success, this study used a composite score of 3 distinct domains, namely pain, function, and satisfaction to define success. Although the individual components could independently serve as an external anchor, the combination of 3 such criteria strengthens the validity of the external anchor and represents a more comprehensive and clinically meaningful definition of success. In addition, the threshold was determined using an anchor-based approach, and was corroborated using a distribution-based, 75th centile comparison of the ROC findings. Finally, the ROC curves and the thresholds were found to remain consistent over time in terms of accuracies and observed thresholds at both the 1-year and 2-year marks, as well as by demographic subgroup.

Limitations

This study should be viewed in light of its limitations. It is important to note that this study pertains only to the THA population and the results reported presently cannot be transferred to other populations previously evaluated by the FJS such as total knee arthroplasty or anterior cruciate ligament reconstruction patients.

Although using a composite score as an external criterion is in our opinion a strength of this study, arguments can be made to support other methods of assessment. In light of this understanding, this study used tools and cutoff points which were supported by literature. The International Society of Arthroplasty Registries guideline recommendation for pain assessment was used for the pain benchmark. The benchmark for satisfaction was based on previous studies with a similar methodology. Although no consistent recommendation or method has been used in previous studies for function, this study used the functional components of the HHS—the most commonly reported and collected PRO in the field of hip arthroplasty [35]. Although the HHS is an accepted tool in assessment of hip pathologies and treatments, it is recognized that this was an arbitrary decision, and other determinants of function may result in different findings. Additionally, despite our attempt to base cutoff points for success on previous literature, other criteria could be suggested.

A second limitation of this study was the retrospective nature of the study, and the requirement to have complete follow-up data at both 1 and 2 years may introduce selection and loss-to-follow-up bias.

A third limitation of this study is the inclusion of a heterogenous population of patients, undergoing various approaches of THA, including posterior and anterior approaches, as well as robotic-assisted and fluoroscopic-guided component placement. However, the goal was to determine a corresponding score for success, based on the patient's perception in the postoperative course, regardless of which methods were utilized during surgery. Additionally, by including a heterogenous population with various approaches and guidance methods, this study is more generalizable to the general THA population which also includes a similarly heterogenous population. An attempt to evaluate the effect of heterogeneity was made by evaluating the ROC curves and thresholds based on age, gender and BMI, which did not find any significant differences.

A fourth limitation of this study was that no preoperative FJS assessment was made. As the FJS was designed to evaluate an artificial joint, the authors did not include it in the preoperative patient assessment. Although a general form applicable to native joints has been designed, it has not been implemented at our institution.

Finally, although the heterogenicity of the population contributes to generalizability of the results, the patients were all operated on at a single institution, by a single surgeon, a fact which may reduce the generalizability of the results. Additionally, only patients with complete follow-up at both 1 and 2 years following THA were included, which may produce a sampling bias.

In conclusion, The FJS has excellent accuracy in demonstrating successful outcome following THA. The FJS threshold for success at 1 and 2 years postoperatively is 73.96 and 69.79, respectively. The higher rates of success at 2 years, along with a rise in the mean FJS, may indicate continued clinical improvement up to 2 years after THA.

References

[1] Collins NJ, Roos EM. Patient-reported outcomes for total hip and knee arthroplasty. Clin Geriatr Med 2012;28:367–94. https://doi.org/10.1016/j.cger.2012.05.007.

[2] Thienpont E, Opsomer G, Koninckx A, Houssiau F. Joint awareness in different types of knee arthroplasty evaluated with the Forgotten Joint score. J Arthroplasty 2014;29:48–51. https://doi.org/10.1016/j.arth.2013.04.024.

[3] 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;27:430–436.e1. https://doi.org/10.1016/j.arth.2011.06.035.

[4] Hamilton DF, Loth FL, Giesinger JM, Giesinger K, MacDonald DJ, Patton JT, 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:218–24. https://doi.org/10.1302/0301-620X.99B2.BJJ-2016-0606.R1.

[5] Ferreira MC, Silva G, Zidan FF, Franciozi CE, Luzo MVM, Abdalla RJ. Forgotten Joint Score - Portuguese translation and cultural adaptation of the instrument of evaluation for hip and knee arthroplasties. Rev Bras Ortop 2018;53:221–5. https://doi.org/10.1016/j.rboe.2018.02.006.

[6] Klouche S, Giesinger JM, Sariali E-H. Translation, cross-cultural adaption and validation of the French version of the Forgotten Joint Score in total hip arthroplasty. Orthop Traumatol Surg Res 2018;104:657–61. https://doi.org/10.1016/j.otsr.2018.04.010.

[7] Baumann F, Ernstberger T, Loibl M, Zeman F, Nerlich M, Tibesku C. Validation of the German Forgotten Joint Score (G-FJS) according to the COSMIN checklist: does a reduction in joint awareness indicate clinical improvement after arthroplasty of the knee? Arch Orthop Trauma Surg 2016;136:257–64. https://doi.org/10.1007/s00402-015-2372-x.

[8] 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;82:112–8.

[9] Cao S, Liu N, Han W, Zi Y, Peng F, Li L, 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 2017;12:6. https://doi.org/10.1186/s13018-016-0508-5.

[10] Matsumoto M, Baba T, Homma Y, Kobayashi H, Ochi H, Yuasa T, et al. Validation study of the Forgotten Joint Score-12 as a universal patient-reported outcome measure. Eur J Orthop Surg Traumatol 2015;25:1141–5. https://doi.org/10.1007/s00590-015-1660-z.

[11] Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. Ann Rheum Dis 2005;64:34–7. https://doi.org/10.1136/ard.2004.023028.

[12] Jevsevar DS, Sanders J, Bozic KJ, Brown GA. An introduction to clinical significance in orthopaedic outcomes research. JBJS Rev 2015;3. https://doi.org/10.2106/JBJS.RVW.N.00064.

[13] Hamilton DF, Loth FL, MacDonald DJ, Giesinger K, Patton JT, Simpson AH, et al. Treatment success following joint arthroplasty: defining thresholds for the Oxford hip and knee scores. J Arthroplasty 2018;33:2392–7. https://doi.org/10.1016/j.arth.2018.03.062.

[14] Rolfson O, Eresian Chenok K, Bohm E, Lübbeke A, Denissen G, Dunn J, et al. Patient-reported outcome measures in arthroplasty registries. Acta Orthop 2016;87(Suppl. 1):3–8. https://doi.org/10.1080/17453674.2016.1181815.

[15] Judge A, Arden NK, Kiran A, Price A, Javaid MK, Beard D, et al. Interpretation of patient-reported outcomes for hip and knee replacement surgery. J Bone Joint Surg Br 2012;94-B:412–8. https://doi.org/10.1302/0301-620X.94B3.27425.

[16] Keurentjes JC, Van Tol FR, Fiocco M, So-Osman C, Onstenk R, Koopman-Van Gemert AW, et al. Patient acceptable symptom states after total hip or knee replacement at mid-term follow-up: thresholds of the Oxford hip and knee scores. Bone Joint Res 2014;3:7–13. https://doi.org/10.1302/2046-3758.31.2000141.

[17] 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;90:253–7. https://doi.org/10.1080/17453674.2019.1599252.

[18] Hamilton DF, Giesinger JM, MacDonald DJ, Simpson AH, Howie CR, Giesinger K. Responsiveness and ceiling effects of the Forgotten Joint Score-12 following total hip arthroplasty. Bone Joint Res 2016;5:87–91. https://doi.org/10.1302/2046-3758.53.2000480.

[19] Thompson SM, Salmon LJ, Webb JM, Pinczewski LA, Roe JP. Construct validity and test re-test reliability of the Forgotten Joint Score. J Arthroplasty 2015;30:1902–5. https://doi.org/10.1016/j.arth.2015.05.001.

[20] Giesinger JM, Hamilton DF, Jost B, Behrend H, Giesinger K. WOMAC, EQ-5D and Knee Society Score thresholds for treatment success after total knee arthroplasty. J Arthroplasty 2015;30:2154–8. https://doi.org/10.1016/j.arth.2015.06.012.

[21] Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Ann Rheum Dis 2005;64:29–33. https://doi.org/10.1136/ard.2004.022905.

[22] Salaffi F, Carotti M, Gutierrez M, Di Carlo M, De Angelis R. Patient acceptable symptom state in self-report questionnaires and composite clinical disease index for assessing rheumatoid arthritis activity: identification of cut-off points for routine care. Biomed Res Int 2015;2015:930756. https://doi.org/10.1155/2015/930756.

[23] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–45.

[24] Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36. https://doi.org/10.1148/radiology.143.1.7063747.

[25] Cohen J. Statistical power analysis for the behavioral sciences. New York: Erlbaum Associates; 1988.

[26] Giesinger JM, Behrend H, Hamilton DF, Kuster MS, Giesinger K. Normative values for the Forgotten Joint Score-12 for the US general population. J Arthroplasty 2019;34:650–5. https://doi.org/10.1016/j.arth.2018.12.011.

[27] Puliero B, Blakeney WG, Beaulieu Y, Vendittoli P-A. Joint perception after total hip arthroplasty and the forgotten joint. J Arthroplasty 2019;34:65–70. https://doi.org/10.1016/j.arth.2018.09.086.

[28] MacKay C, Clements N, Wong R, Davis AM. A systematic review of estimates of the minimal clinically important difference and patient acceptable symptom state of the Western Ontario and McMaster Universities Osteoarthritis Index in patients who underwent total hip and total knee replacement. Osteoarthritis Cartilage 2019. https://doi.org/10.1016/j.joca.2019.05.002.

[29] Arden NK, Kiran A, Judge A, Biant LC, Javaid MK, Murray DW, et al. What is a good patient reported outcome after total hip replacement? Osteoarthritis Cartilage 2011;19:155–62. https://doi.org/10.1016/j.joca.2010.10.004.

[30] Browne JP, Bastaki H, Dawson J. What is the optimal time point to assess patient-reported recovery after hip and knee replacement? A systematic review and analysis of routinely reported outcome data from the English patient-reported outcome measures programme. Health Qual Life Outcomes 2013;11:128. https://doi.org/10.1186/1477-7525-11-128.

[31] Naal FD, Impellizzeri FM, Lenze U, Wellauer V, von Eisenhart-Rothe R, Leunig M. Clinical improvement and satisfaction after total joint replacement: a prospective 12-month evaluation on the patients' perspective. Qual Life Res 2015;24:2917–25. https://doi.org/10.1007/s11136-015-1042-3.

[32] Lim CR, Harris K, Dawson J, Beard DJ, Fitzpatrick R, Price AJ. Floor and ceiling effects in the OHS: an analysis of the NHS PROMs data set. BMJ Open 2015;5:e007765. https://doi.org/10.1136/bmjopen-2015-007765.

[33] Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012.

[34] Thienpont E, Berghe AV, Schwab PE, Forthomme JP, Cornu O. Joint awareness in osteoarthritis of the hip and knee evaluated with the 'Forgotten Joint' Score before and after joint replacement. Knee Surg Sports Traumatol Arthrosc 2016;24:3346–51. https://doi.org/10.1007/s00167-015-3970-4.

[35] Siljander MP, McQuivey KS, Fahs AM, Galasso LA, Serdahely KJ, Karadsheh MS. Current trends in patient-reported outcome measures in total joint arthroplasty: a study of 4 major orthopaedic journals. J Arthroplasty 2018;33:3416–21. https://doi.org/10.1016/j.arth.2018.06.034.