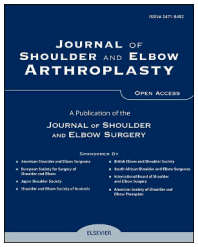
Journal of Shoulder and Elbow Arthroplasty 10 (2026) 100059

Available online at www.sciencedirect.com

ScienceDirect

www.sciencedirect.com/journal/journal-of-shoulder-and-elbow-arthroplasty



The reliability of the forgotten joint score for shoulder arthroplasty

Madilyn L. Feik, BS[a], *Ashley Z. Smith, BS*[a], *Kevin K. Chen, MD*[b], *Zachary A. Gehring, MS*[a], *Natalie L. Myers, PhD*[c], *and James M. Gregory, MD*[a,*]

[a]*Department of Orthopaedic Surgery, The University of Texas Health Science Center at Houston, McGovern Medical School, Houston, TX, USA*

[b]*Hackensack University Medical Center, Hackensack Meridian Health, Hackensack, NJ, USA*

[c]*ATI Physical Therapy, Institute for Musculoskeletal Advancement, Downers Grove, IL, USA*

ABSTRACT

Background: The Forgotten Joint Score, a patient-reported outcome measure assessing the ability to forget the artificial joint, has demonstrated minimal ceiling effects in hip and knee arthroplasty patients. This study determines the reliability of an upper extremity 12-item Forgotten Joint Score (FJS-12) in shoulder arthroplasty patients.

Methods: The FJS-12 was developed from the previously published 70-item list assessing the Forgotten Joint Score concept in shoulder patients. Sixty patients with glenohumeral osteoarthritis completed the FJS-12 at 6 months (n = 38) or 1 year (n = 22) after anatomic or reverse total shoulder arthroplasty and again at least 48 hours later. Test-retest reliability analysis was conducted using intraclass correlation coefficient ($ICC_{2,1}$) with 95% confidence intervals (CIs), standard error of measure (SEM_{90}), and minimal detectable change (MDC_{90}).

Results: FJS-12 test-retest time frame was 8 ± 5 days (range 2-19 days). Six-month FJS-12 was 64 ± 33 with $ICC_{2,1} = 0.97$, 95% CI [0.94-0.98], $SEM_{90} = 9.6$, and $MDC_{90} = 13.5$. One-year FJS-12 was 70 ± 29 with $ICC_{2,1} = 0.96$, 95% CI [0.91-0.98], $SEM_{90} = 10.0$, and $MDC_{90} = 14.2$.

Conclusion: The FJS-12 demonstrated excellent test-retest reliability, representing an important step toward a more sensitive measure to evaluate shoulder arthroplasty outcomes.

Level of evidence: Level III; Prospective Diagnostic Study

Keywords: Patient-reported outcome measure; Forgotten Joint Score; Total shoulder arthroplasty; Reliability; Ceiling effect; FJS

All participants read and signed an electronic informed consent form approved by the University of Texas Health Science Center at Houston (IRB approval #:HSC-MH-21-1041) prior to enrollment in the study.

*Corresponding author: James M. Gregory, MD, FAAOS, FAOA, Department of Orthopaedic Surgery, McGovern Medical School at UTHealth Houston, 950 Corbindale Road Suite 300, Houston, TX 77024, USA.

E-mail address: James.M.Gregory@uth.tmc.edu (J.M. Gregory).

https://doi.org/10.1016/j.jsea.2026.100059

Shoulder arthroplasty is an evolving field; the use of both anatomic shoulder arthroplasty (aTSA) and reverse total shoulder arthroplasty (rTSA) has continued to grow over the last 2 decades.[11,12,34] While range of motion and radiographic imaging are important objective metrics for surgical outcomes,[15,38] patient-reported outcome measures (PROMs)[32] provide insight into treatment effectiveness from the patient's perspective, shifting the focus of outcomes from pathologically-centered to patient-centered.[13,17] When selecting a PROM for clinical practice, consideration should be given to its breadth (disease-specific vs. region-specific vs. generic health-specific), psychometric properties, scoring criteria, and degree of ceiling effects within the population of interest during validity-based testing.[20,25,35]

The Forgotten Joint Score (FJS) is a PROM assessing a patient's ability to forget their artificial joint, where a fully "forgotten" joint represents the highest level of patient satisfaction.[4] The final 12 item Forgotten Joint Score (FJS-12), was initially created to address ceiling effects observed with common PROMs following lower extremity arthroplasty.[4] The ceiling effect concept is a statistical construct that describes the clustering of a patient's score toward the upper (ceiling) limit of a scale which can inhibit the instrument's responsiveness and sensitivity to change over time.[16] Unacceptable ceiling effects are generally regarded as 15% or more of the study population achieving the highest possible score on a given PROM. Outcome measures with limited ceiling effects can detect small differences among patients to better assess the success of the intervention, which is particularly important when assessing higher functioning individuals and successful procedures. Ceiling effects are particularly prevalent when evaluating patients undergoing joint arthroplasty, as the majority of these individuals report successful long-term outcomes following the surgical intervention.[3,4,7,40] As the FJS-12 has demonstrated lower ceiling effects in hip and knee arthroplasty when compared to other readily available PROMs for these populations,[16,49] an FJS questionnaire applicable to the shoulder arthroplasty population may also be less susceptible to ceiling effects than existing shoulder PROMs.

PROMs used to evaluate aTSA and rTSA limited by ceiling effects include the American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form, Simple Shoulder Test (SST), University of California Los Angeles (UCLA), Oxford Shoulder Score, and Shoulder Pain and Disability Index (SPADI).[7,10,44,46] PROMs that are not joint- or condition-specific, such as the Disabilities of the Arm, Shoulder, and Hand and Patient-Reported Outcomes Measurement Information System, are also susceptible to ceiling effects because they assess broader domains that may not be relevant to the shoulder arthroplasty population.[39] Although less ceiling effects are observed with the Constant-Murley Shoulder Outcome Score, 65% of the measure depends on proper range of motion and strength measurement, yielding limited understanding of outcomes from the patient's perspective.[26] Many other shoulder PROMs also have limitations beyond ceiling effects; notable shortcomings include inconsistent item responses and patient uncertainty or inaccuracy during form completion (as observed with the ASES),[5] large measurement error associated with the SPADI,[51] and the use of multiple questions contributing to a single response on the UCLA, which may lead to confusion and nonspecific responses.[54] Additionally, PROMs commonly used to assess aTSA and rTSA, such as the ASES, UCLA, Oxford Shoulder Score, and SPADI, require patients to indicate the highest level of activity they can perform or to rate the difficulty experienced when performing these activities.[46] However, patients' perception of functional "difficulty" may not necessarily equate to their degree of discomfort or fear of movement during that task. Comprised of a single question evaluating shoulder function as a percentage of normal, the Single Assessment Numeric Evaluation (SANE) remains a popular PROM for its efficiency and ease of administration, although it only provides information about overall patient status and is not specific to shoulder arthroplasty.[14] These limitations highlight the need to develop a comprehensive yet succinct PROM specific to shoulder arthroplasty patients that has limited susceptibility to ceiling effects.

Given the success of the FJS-12 in hip and knee arthroplasty and ceiling effects observed with common PROMs for the shoulder arthroplasty population, there is potential utility for a parallel version of the FJS-12 in aTSA and rTSA patients. Giesinger et al[21] previously developed a 70-item list assessing the forgotten joint concept in patients with shoulder pathology; however, a succinct yet thorough 12-item questionnaire to evaluate shoulder patients has yet to be developed. This study evaluates the test-retest reliability of a FJS-12 for individuals with glenohumeral osteoarthritis either 6 months or 1 year following aTSA or rTSA.

Materials and methods

Study design

Data for this study were prospectively collected from January 26, 2024, to February 5, 2025. Data collection occurred at the patient's 6-month or 1-year postoperative follow-up appointment using REDCap (Research Electronic Data Capture; Vanderbilt University, Nashville, TN, USA) electronic data capture tools hosted by UTHealth's School of Biomedical Informatics.[23,24] REDCap is a secure, web-based software platform designed to support data capture for research studies, providing 1) an intuitive interface for validated data capture; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for data integration and interoperability with external sources.[23,24] All patients read and signed an electronic informed consent form approved by The University of Texas Health Science Center at Houston (IRB approval #: HSC-MH-21-1041) prior to enrollment in the study.[31]

Patients

Patients reporting to a single board-certified orthopedic surgeon following routine arthroplasty were recruited. Patients were included in the study if they had a primary diagnosis of glenohumeral osteoarthritis, underwent primary aTSA or rTSA, and had a follow-up time point of 6 months or 1 year from their surgical date. Patients were excluded if they had a

primary diagnosis other than glenohumeral osteoarthritis, underwent revision arthroplasty, or English was not their primary speaking and reading language.

Procedures

The 70-item list by Giesinger et al,[21] developed following comprehensive literature review, international expert board evaluation, and patient feedback, was consolidated into 7 activity-based domains-activities of daily living, reaching up, reaching behind, dressing or undressing, personal hygiene, sitting or resting, and carrying weight. To decrease respondent burden and remove redundancy, a total of 12 questions, representing each domain, were selected based on 2 key factors: 1) including an appropriate spectrum of activity and 2) emphasizing relevant functions specific to the TSA population. The study design was communicated to the owners of the Forgotten Joint Score, and a user agreement was obtained. Once the 12 questions were identified, patients meeting inclusion criteria were asked to provide demographic information and complete the FJS-12 at 2 different time points. During time point one (TP_{one}), patients were consented and asked to complete this information in one of 2 ways: 1) in person on an iPad or 2) over the phone with a member of the research team sending the patient information via a REDCap secure link. TP_{one} was defined as the patients 6-month or 1-year postoperative appointment. The follow-up phone call for time point 2 (TP_{two}) was at least 48 hours following TP_{one}. If the patient did not answer on the first attempt a voicemail was left and a second call was initiated one week later. Therefore, a range from 48 hours to 21 days was accepted in order to accommodate patients' schedules. If data were collected from a patient at the 6-month postoperative visit, it was not collected again at the 1-year postoperative visit.

Patients were required to answer all 12 questions on the FJS-12 (Table I). FJS-12 scoring is synonymous to the original FJS-12 instituted on knee and hip arthroplasties, with a total score ranging from 0-100.[4] Responses to each of the 12 questions are rated on a 0-4 point Likert scale, summed, and divided by the number of completed items to obtain a mean score. This value is then multiplied by 25 and subtracted from 100, changing the direction of the final score so that higher scores indicate a higher degree of forgetting the artificial joint.[4]

Statistical analysis

A priori power analysis was conducted to determine the sample size required for assessing reliability using the intraclass correlation coefficient (ICC). The analysis was based on detecting an expected reliability of 0.90 compared with a minimum acceptable reliability of 0.75, using a two-tailed test with a significance level (α) of 0.05 and power ($1-\beta$) of 0.80. Assuming 2 repeated measurements per subject ($k = 2$), the required sample size was calculated to be 33 subjects. To account for an anticipated 40% dropout rate, the final total target sample size was adjusted to 55 subjects.

Data were reported as mean ± standard deviation for continuous data and total count (percentage) for categorical data. The test-retest reliability of the FJS-12 was determined using the $ICC_{2,1}$. The ICC values range between 0 and 1; values < 0.5 indicate poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values > 0.9 indicate excellent reliability.[28] Confidence intervals (CIs) were calculated at the 95% level for the reliability coefficients. The error associated with the FJS-12 was calculated by the use of the standard error of measure (SEM) using the following formula: $SEM = SDx[\sqrt{1 - ICC}]$. The SEM was then multiplied by the z-score (1.64 for 90% CI) for the true score about the observed score (SEM_{90}). To determine the smallest amount of change that can be detected within the FJS-12, the minimal detectable change (MDC) was calculated: $MDC = SEMx[\sqrt{2}]$. The MDC was then multiplied by the z-score (1.64 for 90% CI), indicating a real change beyond measurement error with 90% confidence (MDC_{90}). These reliability statistics were employed on the entire cohort, 6-month postoperative cohort, and 1-year postoperative cohort.

Results

Within the time frame of the study (January 26, 2024 – February 5, 2025), 168 patients were available for follow-up based on aTSA and rTSA time frames. Seventy-one of these patients did not meet inclusion criteria, as 31 had undergone revision arthroplasty and 40 had a diagnosis other than osteoarthritis. Of the remaining 97 patients available for enrollment, 21 were unable to be reached for consent, 9 never completed TP_{two}, and 7 were outside the 21-day follow-up window (minimum 23 days – maximum 63 days) that the authors deemed acceptable for TP_{two} data collection. This left 60 patients to be included in the final sample for analysis.

Patient demographics can be found in Table II. The cohort was 42% female and 58% male with a mean age of 68 ± 11 years. The right shoulder was injured in 52% of patients. The average time between the 2 follow-up time points (test-retest time frame) was 8 ± 5 days, with a minimum of 2 days and maximum of 19 days. Effort was made to contact eligible patients on the exact 6-month and 1-year postoperative anniversary date; however, there was some variation in the number of days, ranging from 0-101 days. The average days equated to within 22 ± 27 days.

The average FJS-12 score for the 6-month postoperative cohort at TP_{one} was 64 ± 33. In the 1-year postoperative cohort, the average FJS-12 score at TP_{one} was 70 ± 29. Reliability statistics for the FJS-12 are reported in Table III. $ICC_{2,1}$ (95% CI) values were 0.96 (0.95-0.97) in the entire cohort, 0.97 (0.94-0.98) at 6 months postoperative, and 0.96 (0.91-0.98) at 1-year postoperative, indicating excellent test-retest reliability based on each $ICC_{2,1}$ and its lower bound of the 95% CI exceeding 0.90. In the 6 months postoperative cohort, SEM was 5.85 (SEM_{90}, 9.60) and MDC was 8.2 (MDC_{90}, 13.5). The 1-year postoperative cohort demonstrated a SEM of 6.15 (SEM_{90}, 10.09) and an MDC of 8.6 (MDC_{90}, 14.2).

Table I – The Forgotten Joint Score 12-item questionnaire.

The following 12 questions refer to how aware you are of your shoulder or artificial shoulder joint in everyday life. Please select one answer from each question. Are you aware of your shoulder or artificial shoulder joint…
…before falling asleep?
…taking a shower?
…taking off a coat or jacket?
…during do it yourself (DIY) jobs around the house?
…with light recreational activity?
…reaching overhead to a high shelf?
…performing a sudden movement?
…getting out of a car?
…scratching between shoulder blades?
…wiping bottom?
…carrying a shopping bag?
…sitting for an hour?
never = 0; almost never = 1; seldom = 2; sometimes = 3; mostly = 4.

Table II – Patient characteristics presented as mean ± standard deviation or total count (percentage) data for all 3 cohorts analyzed.

Demographic data	Entire cohort (n = 60)	6 mo postoperative (n = 38)	1 y postoperative (n = 22)
Age (years)	68 ± 11	68 ± 11	68 ± 8
Height (cm)	169 ± 11	167 ± 12	173 ± 10
Weight (kg)	85 ± 20	83 ± 21	88 ± 17
BMI (kg/m^2)	29 ± 7	29 ± 7	29 ± 6
Previous shoulder surgery			
Yes	12 (20%)	10 (26%)	2 (9%)
No	48 (80%)	28 (74%)	20 (91%)
Sex			
Female	25 (42%)	17 (45%)	8 (37%)
Male	35 (58%)	21 (55%)	14 (63%)
Race			
Caucasian	50 (83%)	32 (84%)	18 (82%)
Other	10 (17%)	6 (16%)	4 (18%)
Arm dominance*			
Right	39 (81%)	24 (83%)	15 (79%)
Left	9 (19%)	5 (17%)	4 (21%)
Injured arm			
Right	31 (52%)	23 (61%)	8 (36%)
Left	29 (48%)	15 (39%)	14 (64%)
Arthroplasty type			
aTSA	49 (82%)	33 (87%)	16 (73%)
rTSA	11 (18%)	5 (13%)	6 (27%)

aTSA, anatomic total shoulder arthroplasty; *BMI*, body mass index; *rTSA*, reverse total shoulder arthroplasty.
* Missing data on 12 patients for entire cohort, 9 patients for 6-mo postoperative, and 3 patients for 1-y postoperative.

Discussion

The FJS-12, a consolidation of the systematically developed and rigorously evaluated 70-item question pool into a practical, clinically relevant 12-item questionnaire, demonstrated excellent test-retest reliability across all 3 cohorts in this study. Our findings suggest that the FJS-12 can be used consistently in patients undergoing aTSA or rTSA at 2 different time points. As such, it is reasonable to believe that these reliability findings would be similar if collected at a different time point (ie, baseline FJS-12 scores prior to surgery).

The original lower extremity FJS-12 was developed because as outcomes of joint replacement continue to improve over time, ceiling effects increasingly limit differentiation between successful outcomes.[16] While the growing usage of the original FJS-12 in hip, knee, and ankle populations reflects the need for a more sensitive PROM in lower extremity arthroplasty, ceiling effects of common shoulder PROMs increasingly limit interpretation of successful outcomes in upper extremity arthroplasty as well. At a minimum 2-year follow-up, unacceptable ceiling effects were reported in aTSA patients for the SST (43.4%), ASES (23.7%), and UCLA (22.2%), and in rTSA patients for the SST (34.1%).[46] Similar findings were observed in another study, with postoperative ceiling effects in aTSA patients of 52.9%, 25.1%, 20.8%, and 25.6% for the SST, ASES, UCLA, and SPADI, respectively, and in rTSA patients of 36.7% for the SST.[44] These findings highlight the need to investigate a

Table III – Reliability statistics for all 3 cohorts.

Reliability statistic	Entire cohort (n = 60)	6-mo postoperative (n = 38)	1-y postoperative (n = 22)
$ICC_{2,1}$ (95% CI)	0.96 (0.95, 0.97)	0.97 (0.94, 0.98)	0.96 (0.91, 0.98)
SEM	6.54	5.85	6.15
SEM_{90}	10.73	9.60	10.09
MDC	9.2	8.2	8.6
MDC_{90}	15.1	13.5	14.2

CI, confidence interval; *ICC*, intraclass correlation coefficient; *MDC*, minimal detectable change; *SEM*, standard error of measure.

parallel version of the FJS-12 in the shoulder arthroplasty population.

The results of this study are similar to other studies that have used the FJS-12 as a PROM for lower joint arthroplasties. Compared to the total knee arthroplasty (TKA) population, our study found slightly higher 6-month FJS-12 scores with lower 1-year scores (mean scores after TKA were 59 ± 28 and 73 ± 24 at 6 months and 1 year, respectively).[8] Psychometric properties of the FJS-12 among TKA patients were found to be 5.2 and 12 points for the SEM and MDC_{90}, respectively[9]; these properties are similar to the findings of our study. Much lower FJS-12 scores have been documented at the 6-month time point in the total hip arthroplasty (THA) population (50 ± 30)[42] with scores comparable to our study at 1-year (70 ± 29).[45] Lower FJS-12 scores were observed at 1-year when compared to the total ankle replacement (64 ± 20) and ankle arthrodesis (58 ± 26) populations.[37] Six-month data in the ankle population have not yet been reported in the literature. The slightly higher scores observed in our study at both time points when compared to the scores for TKA, THA, total ankle replacement, and ankle arthrodesis may relate to the weightbearing status of these lower extremity joints as compared to the shoulder. Additionally, patients undergoing TSA are reported to have fewer complications, shorter length of stay, and lower total charges than those who undergo TKA or THA, which may contribute to the higher self-reported outcome scores.[18]

When comparing the FJS-12 results of this study to the commonly implemented American Shoulder and Elbow Surgeons PROM in the TSA population, findings are interesting. Matar et al[33] found patients with glenohumeral osteoarthritis undergoing rTSA had 6-month postoperative ASES scores of 68.77 ± 17.02 and 70.93 ± 17.78 at 1-year; in this study, slightly lower average FJS-12 scores were observed at both the 6-month and 1-year time points. Patients completing the ASES at both the 1-year and 2-year time points have nearly perfect scores ranging between 93 and 95, respectively, following aTSA.[41] Patients undergoing rTSA show similar results on the ASES postoperatively, with most patients scoring above 80 points.[19,47] Furthermore, Kirsch et al[27] reported 2-year postoperative mean ASES scores among glenohumeral osteoarthritis patients treated with primary aTSA and rTSA of 95 and 93.3, respectively. These results indicate that continued room for patient improvement is minimal and ceiling effects may increase, meaning each patient's true outcome cannot be measured appropriately.[53] Consequently, alternative PROMs to evaluate shoulder arthroplasty should be investigated.

Despite the plethora of PROMs available to assess shoulder pathology, a single PROM has not emerged as superior for a given shoulder pathology. Choosing an optimal set that provides a comprehensive assessment of the population of interest can be challenging, and as several questionnaires are often administered concurrently, there is potential for redundancy and unnecessary cost and time burden on both patients and administrators. Sutton et al[48] evaluated compliance and resource utilization for common PROMs used for THA and TKA populations and demonstrated significant institutional cost and time burden along with low compliance rates. The FJS-12 concept can be an attractive option for busy clinicians and researchers assessing arthroplasty patients, given that "forgetting" the joint implies the absence of a variety of domains such as pain, stiffness, discomfort, and fear of movement, eliminating the need to inquire about each of these domains individually.[4] The FJS-12 may similarly be a valuable addition or alternative to existing PROMs by assessing multiple clinical manifestations within a short set of questions.[22,50] Moreover, the ASES, one of the most commonly used PROMs in the assessment of aTSA and rTSA patients,[36] demonstrates strong to excellent correlations with other global shoulder PROMs such as the Constant, SPADI, and SST.[2] These findings suggest that concurrent administration of broadly applicable PROMs may be redundant and contribute to unnecessary survey burden. Contrarily, Descamps and Boileau[14] proposed implementing the FJS concept into the single-question SANE by evaluating shoulder function as a percentage of a fully "forgotten" joint. However, this has not yet been studied in shoulder arthroplasty patients, and the lack of detail obtained with the SANE prohibits contextualization of poor results.[14] Therefore, the FJS-12 can potentially provide the balance between oversimplification and unnecessary burden, serving as a concise yet comprehensive instrument for shoulder arthroplasty.

Although the original FJS-12 was created to assess lower extremity arthroplasty,[4] it has since been expanded to other hip and knee pathologies, with high reliability among patients undergoing arthroscopic treatment of anterior cruciate ligament tears and femoroacetabular impingement.[6,30,43,52] While the FJS-12 was only evaluated in patients with glenohumeral osteoarthritis undergoing TSA in the present study, the FJS-12 could potentially assess patients with rotator cuff arthropathy undergoing TSA or expand to other shoulder populations. However, it is not recommended to interpret FJS-12 and thus FJS-12 results across overlapping pathologies or procedures, as discriminatory power and other psychometric properties

vary based on age and gender in the lower extremity arthroplasty population.[1,29] Furthermore, psychometric property reporting is inconsistent across existing studies evaluating the FJS.[29] Therefore, caution must be taken when using the FJS-12 in broader patient populations, and rigorous evaluation of its psychometric properties within those specific contexts is necessary.

This study is not without limitations that should be acknowledged when interpreting the results. First, follow-up FJS-12 questionnaires were not always conducted precisely at the intended 6-month or 1-year time intervals. Instead, there was a range in the number of days between the target follow-up date and the actual clinic visit, with some assessments occurring exactly on schedule and others up to 101 days later. While every effort was made to collect data as close as possible to the scheduled follow-up, logistical constraints made consistency difficult and resulted in time point collection variation. However, this reflects real-world clinical practice, where perfect adherence to follow-up schedules is often not feasible. While patients who missed in-person appointments were attempted to be contacted by phone, data collection often depended on patients attending their in-person follow-up visits. This may have introduced selection bias, as individuals more engaged in their postoperative care were more likely to be included, and patients choosing not to complete the TP_{two} questionnaire were excluded from reliability analysis. Additionally, patients may have completed the TP_{one} questionnaire in person while the TP_{two} questionnaire was completed over the phone. A potential limitation of this approach is that varying levels of technological literacy among patients influenced participation, and differing methods of data collection (reading vs. listening) may have influenced patients' comprehension and responses. Furthermore, while the study aimed to collect TP_{two} at 48 hours, the average time between TP_{one} and TP_{two} was eight days. Despite these limitations, our study mirrors the practical challenges faced in routine clinical follow-up and provides valuable insight into the psychometric properties of the FJS-12 across a range of real-world scenarios. Additionally, certain items within the FJS-12 may be overly broad, particularly the question regarding "light recreational activity." This nonspecific question was intended to capture patient-perceived light activity and address limitations of existing PROMs in assessing patients returning to higher activity levels following shoulder arthroplasty; however, because patients may interpret this term to include activities of different intensities (eg, walking, swimming, or golf), the variability in patient interpretation could reduce the content validity of this item. Future refinements of the questionnaire may benefit from clarifying or stratifying this activity level to enhance comparability across patient groups. Future research is underway to further validate the FJS-12 used in this study, and subsequent studies should evaluate the FJS-12 at the 2-year postoperative time point.

Conclusion

A PROM that is sensitive to change with limited susceptibility to ceiling effects is necessary for the shoulder arthroplasty population. As the original FJS-12 has been successful in combatting ceiling effects following hip and knee arthroplasty, this study investigated a shoulder version of the FJS-12 as an option for measuring shoulder arthroplasty outcomes, finding excellent test-retest reliability between responses at different time points. Our research group is currently working to validate this score with the hope to eventually improve understanding of patients' joint awareness following shoulder arthroplasty.

Disclaimers:

Funding: No funding was disclosed by the authors.

Conflicts of interest: James M. Gregory reports that he has received Arthrex research support and consulting payments, Stryker research support and consulting payments, and Sparta Biomedical stock/stock options. Any additional authors, their immediate families, and any research foundations with which they are affiliated have not received any financial payments or other benefits from any commercial entity related to the subject of this article.

REFERENCES

1. Alfaro-Adrián J. Editorial commentary: the forgotten joint Score is valid, but should not be used to compare heterogeneous procedures. Arthroscopy 2023;39:2096–7. https://doi.org/10.1016/j.arthro.2023.03.005.
2. Angst F, Pap G, Mannion AF, Herren DB, Aeschlimann A, Schwyzer HK, et al. Comprehensive assessment of clinical outcome and quality of life after total shoulder arthroplasty: usefulness and validity of subjective outcome measures. Arthritis Rheum 2004;51:819–28. https://doi.org/10.1002/art.20688.
3. Barry LW, Katayama ES, Barnett JS, Henderson BL, Patel AV, Cvetanovich GL, et al. Functionality, complications, and survivorship of total shoulder arthroplasty in patients under 60 years old. J Orthop 2024;55:59–63. https://doi.org/10.1016/j.jor.2024.04.007.
4. 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 Arthoplasty 2012;27:430–436.e1. https://doi.org/10.1016/j.arth.2011.06.035.
5. Bourget-Murray J, Frederick A, Murphy L, French J, Barwood S, LeBlanc J. Establishing user error on the patient-reported component of the American shoulder and elbow surgeons shoulder score. Orthop J Sports Med 2020;8:2325967120910094. https://doi.org/10.1177/2325967120910094.
6. Bramming IB, Kierkegaard S, Lund B, Jakobsen SS, Mechlenburg I. High relative reliability and responsiveness of the forgotten joint score-12 in patients with femoroacetabular impingement undergoing hip arthroscopic treatment. A prospective survey-based study. J Hip Preserv Surg 2019;6:149–56. https://doi.org/10.1093/jhps/hnz020.
7. Brochin RL, Zastrow RK, Patel AV, Parsons BO, Galatz LM, Flatow EL, et al. Long-term clinical and radiographic outcomes of total shoulder arthroplasty in patients under age 60 years. J Shoulder Elbow Surg 2022;31:S63–70. https://doi.org/10.1016/j.jse.2022.02.002.
8. 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 Arthoplasty 2018;33:700–3. https://doi.org/10.1016/j.arth.2017.09.063.

9. Clement ND, Scott CE, Hamilton DF, MacDonald D, Howie CR. Meaningful values in the Forgotten Joint Score after total knee arthroplasty: minimal clinical important difference, minimal important and detectable changes, and patient-acceptable symptom state. Bone Joint J 2021;103:846–54. https://doi.org/10.1302/0301-620X.103B5.BJJ-2020-0396.R1.
10. Davies AR, Sabharwal S, Liddle AD, Zamora-Talaya B, Rangan A, Reilly P. Patient-reported outcomes following total shoulder arthroplasty and hemiarthroplasty: an analysis of data from the National Joint Registry. J Shoulder Elbow Surg 2024;33:2411–20. https://doi.org/10.1016/j.jse.2024.01.053.
11. Day JS, Lau E, Ong KL, Williams GR, Ramsey ML, Kurtz SM. Prevalence and projections of total shoulder and elbow arthroplasty in the United States to 2015. J Shoulder Elbow Surg 2010;19:1115–20. https://doi.org/10.1016/j.jse.2010.02.009.
12. Day JS, Paxton ES, Lau E, Gordon VA, Abboud JA, Williams GR. Use of reverse total shoulder arthroplasty in the Medicare population. J Shoulder Elbow Surg 2015;24:766–72. https://doi.org/10.1016/j.jse.2014.12.023.
13. De Boer D, Delnoij D, Rademakers J. The importance of patient-centered care for various patient groups. Patient Edu Couns 2013;90:405–10. https://doi.org/10.1016/j.pec.2011.10.002.
14. Descamps J, Boileau P. Postoperative shoulder function assessment: proposing the concept of "Forgotten Shoulder" Over "Normal Shoulder" as the new standard in single assessment Numerical evaluation. Arthroscopy 2024;40:1392–3. https://doi.org/10.1016/j.arthro.2024.02.015.
15. Durchholz H, Salomonsson B, Moroder P, Lambert S, Page R, Audigé L, et al. Core set of radiographic parameters for shoulder arthroplasty monitoring: criteria defined by an international Delphi consensus process. JB JS Open Access 2019;4:e0025. https://doi.org/10.2106/JBJS.OA.19.00025.
16. Eckhard L, Munir S, Wood D, Talbot S, Brighton R, Walter B, et al. The ceiling effects of patient reported outcome measures for total knee arthroplasty. Orthop Traumatol Surg Res 2021;107:102758. https://doi.org/10.1016/j.otsr.2020.102758.
17. Epstein RM, Street RL. The values and value of patient-centered care. Ann Fam Med 2011;9:100–3. https://doi.org/10.1370/afm.1239.
18. Farmer KW, Hammond JW, Queale WS, Keyurapan E, McFarland EG. Shoulder arthroplasty versus hip and knee arthroplasties: a comparison of outcomes. Clin Orthop Relat Res 2007;455:183–9. https://doi.org/10.1097/01.blo.0000238839.26423.8d.
19. Flurin P-H, Marczuk Y, Janout M, Wright TW, Zuckerman J, Roche CP. Comparison of outcomes using anatomic and reverse total shoulder arthroplasty. Bull Hosp Jt Dis 2013;71:S101.
20. Gagnier JJ. Patient reported outcomes in orthopaedics. J Orthop Res 2017;35:2098–108. https://doi.org/10.1002/jor.23604.
21. Giesinger JM, Kesterke N, Hamilton DF, Holzner B, Jost B, Giesinger K. Development of an item list to assess the forgotten joint concept in shoulder patients. BMC Musculoskelet Disord 2015;16:1–8. https://doi.org/10.1186/s12891-015-0520-7.
22. Hamilton D, Giesinger J, MacDonald D, Simpson A, Howie C, 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.
23. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019;95:103208. https://doi.org/10.1016/j.jbi.2019.103208.
24. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42:377–81. https://doi.org/10.1016/j.jbi.2008.08.010.
25. Johnston BC, Patrick DL, Devji T, Maxwell LJ, Bingham IIICO, Beaton DE, et al. Patient-reported outcomes. Available at: https://www.cochrane.org/authors/handbooks-and-manuals/handbook/current/chapter-18#section-18-3-1; 2019. Accessed May 1, 2015.
26. King JJ, Gil JN, Wright TW, Vasilopoulos T, Struk AM, Hoover A, et al. Correlation and conversion of the normalized constant Score and patient-reported outcomes in shoulder arthroplasty. J Am Acad Orthop Surg Glob Res Rev 2023;7:e22. https://doi.org/10.5435/JAAOSGlobal-D-22-00247.
27. Kirsch JM, Puzzitiello RN, Swanson D, Le K, Hart P-A, Churchill R, et al. Outcomes after anatomic and reverse shoulder arthroplasty for the treatment of glenohumeral osteoarthritis: a propensity score-matched analysis. J Bone Joint Surg Am 2022;104:1362–9. https://doi.org/10.2106/JBJS.21.00982.
28. Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155–63. https://doi.org/10.1016/j.jcm.2016.02.012.
29. Kuhns BD, Harris WT, Domb BG. Low ceiling effects of the forgotten joint Score compared with legacy measures after joint-preserving procedures: a systematic review. Arthroscopy 2023;39:2086–95. https://doi.org/10.1016/j.arthro.2023.01.107.
30. Ladurner A, Giesinger K, Zdravkovic V, Behrend H. The Forgotten Joint Score-12 as a valuable patient-reported outcome measure for patients after first-time patellar dislocation. Knee 2020;27:406–13. https://doi.org/10.1016/j.knee.2019.12.004.
31. Lawrence CE, Dunkle L, McEver M, Israel T, Taylor R, Chiriboga G, et al. A REDCap-based model for electronic consent (econsent): moving toward a more personalized consent. J Clin Transl Sci 2020;3:345–53. https://doi.org/10.1017/cts.2020.30.
32. Makhni EC, Hennekes ME, Baumhauer JF, Muh SJ, Spindler K. AOA critical issues: patient-reported outcome measures: why every orthopaedic practice should be collecting them. J Bone Joint Surg Am 2023;105:641–8. https://doi.org/10.2106/JBJS.22.00587.
33. Matar RN, Gardner TJ, Kassam F, Grawe BM. When do patients truly reach maximal medical improvement after undergoing reverse shoulder arthroplasty? The incidence and clinical significance of pain and patient-reported outcome measure improvement. JSES Int 2020;4:675–9. https://doi.org/10.1016/j.jseint.2020.03.010.
34. Mayfield CK, Korber SS, Hwang NM, Bolia IK, Gamradt SC, Weber AE, et al. Volume, indications, and number of surgeons performing reverse total shoulder arthroplasty continue to expand: a nationwide cohort analysis from 2016-2020. JSES Int 2023;7:827–34. https://doi.org/10.1016/j.jseint.2023.05.002.
35. McClimans LM, Browne J. Choosing a patient-reported outcome measure. Theor Med Bioeth 2011;32:47–60. https://doi.org/10.1007/s11017-010-9163-8.
36. Narulla R, Song M, Karunaratne S, Smithers C, Petchell J. Trends in shoulder surgery patient-reported outcome measures. JSES Int 2023;7:653–61. https://doi.org/10.1016/j.jseint.2023.03.010.
37. Noguchi K, Yamaguchi S, Teramoto A, Amaha K, Kanzaki N, Tanaka H, et al. Reliability and validity of the Forgotten Joint Score-12 for total ankle replacement and ankle arthrodesis. PLoS One 2023;18:e0286762. https://doi.org/10.1371/journal.pone.0286762.
38. Ode GE, Ling D, Finocchiaro A, Lai EY, Taylor SA, Dines J, et al. Clinical characteristics and patient-reported outcomes of

total shoulder arthroplasty after anterior stabilization: a retrospective matched control study. J Shoulder Elbow Surg 2020;29:S59–66. https://doi.org/10.1016/j.jse.2020.04.003.
39. Overbeek CL, Nota SP, Jayakumar P, Hageman MG, Ring D. The PROMIS physical function correlates with the QuickDASH in patients with upper extremity illness. Clin Orthop Relat Res 2015;473:311–7. https://doi.org/10.1007/s11999-014-3840-2.
40. Patel AV, White CA, Cirino CM, Duey AH, Ranson WA, Gross BD, et al. Mid-term outcomes following reverse total shoulder arthroplasty. Eur J Orthop Surg 2024;34:799–807. https://doi.org/10.1007/s00590-023-03708-6.
41. Patel M, Cogan CJ, Sahoo S, Cannon D, Grewal G, Owings TM, et al. Postoperative patient-reported outcomes and radiographic findings do not significantly change between 1 and 2 years postoperatively after primary anatomic shoulder arthroplasty. J Shoulder Elbow Surg 2025;34:S50–6. https://doi.org/10.1016/j.jse.2025.02.012.
42. Robinson PG, MacDonald DJ, Macpherson GJ, Patton JT, Clement ND. Changes and thresholds in the Forgotten Joint Score after total hip arthroplasty: minimal clinically important difference, minimal important and detectable changes, and patient-acceptable symptom state. Bone Jt J 2021;103:1759–65. https://doi.org/10.1302/0301-620X.103B12.BJJ-2021-0384.R1.
43. Robinson PG, Rankin C, Murray I, Maempel J, Gaston P, Hamilton D. 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;29:1378–84. https://doi.org/10.1007/s00167-020-06138-7.
44. Roche C, Kumar V, Overman S, Simovitch R, Flurin P-H, Wright T, et al. Validation of a machine learning–derived clinical metric to quantify outcomes after total shoulder arthroplasty. J Shoulder Elbow Surg 2021;30:2211–24. https://doi.org/10.1016/j.jse.2021.01.021.
45. Rosinsky PJ, Chen JW, Lall AC, Shapira J, Maldonado DR, Domb BG. Can we help patients forget their joint? Determining a threshold for successful outcome for the forgotten joint score. J Arthoplasty 2020;35:153–9. https://doi.org/10.1016/j.arth.2019.08.014.
46. Schoch BS, King JJ, Fan W, Flurin P-H, Wright TW, Zuckerman JD, et al. Characteristics of anatomic and reverse total shoulder arthroplasty patients who achieve ceiling scores with 3 common patient-reported outcome measures. J Shoulder Elbow Surg 2022;31:1647–57. https://doi.org/10.1016/j.jse.2022.01.142.
47. Shields E, Koueiter DM, Wiater JM. Rate of improvement in outcomes measures after reverse total shoulder arthroplasty: a longitudinal study with 2-year follow-up. J Shoulder Elbow Arthroplasty 2019;3:2471549219861446. https://doi.org/10.1177/2471549219861446.
48. Sutton RM, Baker CM, D'Amore T, Courtney PM, Krueger CA, Austin MS. Poor patient compliance limits the attainability of patient-reported outcome measure completion thresholds for the comprehensive care for joint arthroplasty model. J Arthroplasty 2023;38:S63–8. https://doi.org/10.1016/j.arth.2022.09.026.
49. Thienpont E, Vanden Berghe A, Schwab P-E, 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.
50. 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.
51. Thoomes-de Graaf M, Scholten-Peeters W, Duijn E, Karel Y, De Vet HC, Koes B, et al. The responsiveness and interpretability of the shoulder pain and disability index. J Orthop Sports Phys Ther 2017;47:278–86. https://doi.org/10.2519/jospt.2017.7079.
52. Vermeijden HD, Yang XA, van der List JP, DiFelice GS. Reliable internal consistency and adequate validity of the forgotten joint Score-12 after primary anterior cruciate ligament repair. Arthrosc Sports Med Rehabil 2021;3:e893–900. https://doi.org/10.1016/j.asmr.2021.03.003.
53. Wang L, Zhang Z, McArdle JJ, Salthouse TA. Investigating ceiling effects in longitudinal data analysis. Multivariate Behav Res 2008;43:476–96. https://doi.org/10.1080/00273170802285941.
54. Wylie JD, Beckmann JT, Granger E, Tashjian RZ. Functional outcomes assessment in shoulder surgery. World J Orthop 2014;5:623. https://doi.org/10.5312/wjo.v5.i5.623.