Role of preoperative PROMIS scores in predicting postoperative outcomes and likelihood of achieving MCID following reverse shoulder arthroplasty

Sreten Franovic
Noah A. Kuhlmann
Collin Schlosser
Alexander Pietroski
Asa G. Buchta

See next page for additional authors

Follow this and additional works at: https://scholarlycommons.henryford.com/orthopaedics_articles
Authors
Sreten Franovic, Noah A. Kuhlmann, Collin Schlosser, Alexander Pietroski, Asa G. Buchta, and Stephanie J. Muh
Role of preoperative PROMIS scores in predicting postoperative outcomes and likelihood of achieving MCID following reverse shoulder arthroplasty

Sreten Franovic, Noah Kuhlmann, Collin Schlosser, Alex Pietroski, Asa Gray Buchta, and Stephanie J. Muh*

Henry Ford Health System, Department of Orthopaedic Surgery, 2799 W Grand Blvd, Detroit, MI 48202, USA

ABSTRACT

Background: The patient-reported outcomes measurement information system (PROMIS) has emerged as an efficient and valid outcome measure in various shoulder surgeries. The purpose of this study was to investigate the influence of preoperative PROMIS scores in predicting postoperative PROMIS scores and the likelihood of achieving a minimal clinically important difference (MCID) following primary reverse total shoulder arthroplasty for cuff tear arthropathy. We hypothesize that preoperative PROMIS scores will influence both postoperative PROMIS scores and the probability of achieving MCID.

Methods: 73 patients undergoing reverse shoulder arthroplasty by a board-certified shoulder and elbow surgeon were given three PROMIS CAT forms: PROMIS Upper Extremity Physical Function CAT v2.0 ("PROMIS-UE"), PROMIS Pain Interference v1.1 ("PROMIS-PI"), and PROMIS Depression v1.0 ("PROMIS-D"). PROMIS CAT domain t scores were assessed for significance between both time points using a Paired Samples t test. Minimal clinically important difference (MCID) was calculated using the distribution method and each PROMIS domain was subsequently assessed for its discriminatory ability in predicting postoperative improvement equal to or greater than the MCID through receiver operating characteristic (ROC) curve analysis.

Results: Our cohort consisted of 73 patients (49.3% male) and an average age of 69.7 years (standard deviation, 11.9). Mean follow-up time point was 9.6 months (standard deviation, 5.0) after surgery. Preoperative PROMIS-UE, PROMIS-PI, and PROMIS-D were 29.5 ± 6.2, 63.3 ± 5.4, and 50.1 ± 9.2, respectively. Each domain significantly improved at 10-months, on average, to 40.9 ± 7.8, 51.4 ± 8.5, 42.6 ± 8.1, respectively. Following the distribution-based method for MCID calculation, we found the following MCID values for PROMIS-UE, PROMIS-PI, and PROMIS-D: 3.1, 2.7, and 4.6, respectively. ROC analysis revealed strong predictive ability for PROMIS-UE (AUC = 0.717, p < 0.05), moderate predictive ability for PROMIS-PI (AUC = 0.634, p < 0.05), and excellent predictive ability for PROMIS-D (AUC = 0.864, p < 0.05). Specifically, preoperative cutoff values of <26.0, >70.0, and >52.5 for PROMIS-UE, PROMIS-PI, and PROMIS-D are especially predictive of achieving MCID.

ARTICLE INFO

Keywords:
PROMIS
MCID
Minimal clinically important difference
Clinically significant outcome
Reverse shoulder arthroplasty
RSA
Patient-reported outcomes

Henry Ford Health System Institutional Review Board Committee approved this study (no. 11361).
*Corresponding author at: Division of Hand and Upper Extremity, Department of Orthopaedic Surgery, Henry Ford Health System, West Bloomfield, MI, USA.
E-mail address: smuh1@hfhs.org (S.J. Muh).

https://doi.org/10.1053/j.sart.2020.05.008
1045-4527/© 2020 American Shoulder and Elbow Surgeons. Published by Elsevier Inc. All rights reserved.
Since the initial approval of the modern reverse total shoulder arthroplasty (RSA) by the Federal Drug Administration in 2004, indications, and therefore the incidence, of the surgery has grown quickly [8,24,26]. Studies have displayed that RSA is an effective surgical option for a wide range of glenohumeral pathologies, including cuff tear arthropathy, massive irreparable rotator cuff tears, and proximal humerus fractures [2,4,18]. These studies predominantly selected range of motion measures and legacy patient-reported outcomes (PRO’s), such as the Simple Shoulder Test, Constant score, and the American Shoulder and Elbow Surgeons score as primary outcomes. In 2004, the National Institute of Health introduced the Patient-Reported Outcomes Measurement Information System (PROMIS), a standardized and highly-efficient outcome collection system [9]. The upper extremity domain of PROMIS (PROMIS-UE) has been recently shown to correlate with legacy PRO’s in upper extremity surgery, particularly in patients with rotator cuff disease and glenohumeral arthritis [17,21].

When determining if a surgery was successful, surgeons may look for meaningful improvements in various health measures. One method to demonstrate these improvements is through the minimal clinically important difference (MCID). This measure represents the smallest score change in an outcome measure that reflects a clinically-significant difference, rather than a merely statistically-significant one. MCID values have been elucidated in a variety of orthopedic cohorts, such as: foot and ankle [13], hand [1], knee [7], and even certain shoulder cohorts [5]. While MCID analysis for total shoulder arthroplasty patients currently exists [5], the anatomic differences and clinical presentation of patients warranting reverse shoulder arthroplasty necessitates a distinct examination of a reverse shoulder arthroplasty cohort.

Thus, the purpose of this study was to examine the influence that preoperative PROMIS scores have on postoperative PROMIS scores, and therefore the probability of achieving MCID after primary RSA for cuff tear arthropathy. We hypothesize that preoperative scores will correlate with postoperative scores, and these preoperative scores can provide predictive utility regarding which patients were likely to postoperatively achieve MCID.

1. Methods

This study was approved by our institutional review board prior to onset of data collection. Patients presenting with cuff tear arthropathy that subsequently underwent reverse shoulder arthroplasty by a board-certified shoulder and elbow surgeon were given three PROMIS CAT forms: PROMIS Upper Extremity Physical Function CAT v2.0 (“PROMIS-UE”), PROMIS Pain Interference v1.1 (“PROMIS-PI”), and PROMIS Depression v1.0 (“PROMIS-D”). PROMIS CAT forms were administered via iPad (Apple, Inc., Cupertino, CA, USA) using a secure, web-based platform for recording and storing research data (REDCap, Vanderbilt University, Nashville, TN, USA). Patients were only approached if they were 18 years of age, could communicate in English, presented with cuff tear arthropathy, and elected to undergo RSA. Exclusion criteria were as follows: refusal to complete preoperative and at least one postoperative (>90 days) set of PROMIS CAT forms, undergoing revision surgery prior to first postoperative time point, proximal humerus fracture, and concomitant infection. The 90-day cutoff was chosen, in accordance with previously established methods in the orthopedic literature [5], to reflect the early postoperative time frame where patients were expecting a return to activity. In our practice, the 90-day period is the earliest we suggest a patient follow-up, on an as-needed basis, if recovery is proceeding as expected. If multiple visits were present after the 90-day point, the latest clinical visit was utilized for analysis.

Given the CAT nature of the PROMIS domains used, order, amount, and type of questions asked differed based on each patient’s response. This yields many advantages, such as decreased administration time and burden to the patient. The scores for each domain are normalized to a mean score of 50 and a standard deviation of 10. Furthermore, a higher score in each domain represents more of the measure in question. Thus, a higher score in PROMIS-UE indicates greater upper extremity physical functioning while a higher score in PROMIS-PI indicates greater interference of pain on a patient’s life.

In addition to PROMIS CAT forms, several patient-centric demographics were collected, such as age, sex, ethnicity, smoking status, employment status, and zip code (Table 1).

As previously established [10], zip code was used to estimate median household income (MHI) by cross referencing with a United States Census Bureau website (https://factfinder.census.gov/faces/nav/jsf/pages/community_facts.xhtml?src=bkmk). Electronic medical records were also reviewed for pre- and postoperative range of motion (ROM) data as well as other clinical characteristics, such as: body mass index (BMI), ASA class, and degree of glenoid version. The following ROM values were collected pre- and postoperatively: abduction, forward flexion, and external rotation. Abduction and external rotation (ABER) and abduction and internal rotation (ABIR) were only collected preoperatively due to infrequency of postoperative reporting. All preoperative PROMIS scores and clinical measures were recorded at the same standard preoperative clinic visit.

1.1. Statistical analysis

All available data (N = 73) was analyzed. Descriptive statistics were calculated for all patient demographic and clinical
characteristics (Tables 1 and 2). PROMIS CAT domain t scores were assessed for significance between both time points using a Paired Samples t test. One-way analysis of variance (ANOVA) was conducted to identify any differences among change in PROMIS domain t scores and among patient-centric factors. Pearson correlation coefficients (r) were determined to show both interdomain correlations and correlations between PROMIS domains and ROM data. Correlation coefficients were interpreted as follows: high (>0.70), high-moderate (0.61–0.69), moderate (0.40–0.60), moderate-weak (0.31–0.39), or weak (<0.31) [23].

Minimal clinically important difference (MCID) was calculated using the distribution method, one-half times the standard deviation of the preoperative PROMIS domain [1]. After establishing MCIDs, each PROMIS domain was assessed for its discriminatory ability in predicting postoperative improvement equal to or greater than the MCID through receiver operating characteristic (ROC) curve analysis. Area under the curve (AUC) analysis was used to determine the predictive ability of each domain, using the following cutoffs: 0.61–0.69, moderate predictive ability; 0.70–0.79, strong predictive ability; >0.80, excellent predictive ability [12]. Corresponding p values were computed for testing the hypothesis that the AUC was greater than 0.50, which corresponds to no predictive ability. Prognostic cutoffs were also assessed from the ROC curve coordinates using 95% specificity [6,14]. Finally, a logistic regression multivariate model was also assessed for its ability to predict the achievement of the MCID for each PROMIS CAT domain, which included patient age, sex, BMI, ASA class, and preoperative PROMIS-UE, PROMIS-PI, and PROMIS-D values. AUC analysis was repeated to determine predictive ability and corresponding p values were generated to distinguish between univariate (just preoperative PROMIS score) and multivariate models.

Finally, scatter plots were generated that displayed change in each PROMIS CAT domain in relation to their respective preoperative domain score. MCID values and prognostic cutoffs were delineated to display areas of the graphic that were more likely to achieve MCID. All analyses used a significance level of 5%. SPSS software was used for all statistical analyses (Released 2017. IBM SPSS Statistics for Windows, Version 25.0; IBM Corp., Armonk, NY, USA).

2. Results

Our study was retrospective by nature. By using Current Procedural Terminology (CPT) code 23472, we identified 123 patients that had shoulder arthroplasty and had reached at
least 90-days follow-up. Of these patients, 28 were removed due to undergoing total shoulder arthroplasty, rather than reverse shoulder arthroplasty. From the remaining 95, 22 were removed due to incomplete PROMIS CAT domains leaving 73 patients in our cohort.

Our cohort consisted of 73 patients (49.3% male) and an average age of 69.7 years (standard deviation, 11.9). Mean follow-up time point was 9.6 months (standard deviation, 5.0) after surgery. All 73 implants showed intactness on radiological exams, at latest follow-up, and no revision surgeries were warranted to date. Further demographic and clinical characteristics can be seen in Tables 1 and 2.

Repeated measures ANOVA showed significant impact of time on each PROMIS CAT domain (p < 0.05). Preoperative PROMIS-UE, PROMIS-PI, and PROMIS-D were 29.5 ± 6.2, 63.3 ± 5.4, and 50.1 ± 9.2, respectively. Each domain significantly improved postoperatively to 40.9 ± 7.8, 51.4 ± 8.5, 42.6 ± 8.1, respectively. Similarly, repeated measures ANOVA showed significant increases in both abduction (+29.2 degrees, p < 0.05) and forward flexion (+27.6°, p < 0.05) ROM measures. Following the distribution-based method for MCID calculation, we found the following MCID values for PROMIS-UE, PROMIS-PI, and PROMIS-D: 3.1, 2.7, and 4.6, respectively. A post hoc analysis identified the percentage of patients meeting MCID for each domain. 82.1% of patients met MCID for PROMIS-UE, 85.2% met MCID for PROMIS-PI, and 55.6% met MCID for PROMIS-D, at mean 10-months postoperatively. No significant differences were noted in % achievement of MCID when comparing those with <1 year follow-up to those with at least 2 year follow-up.

ROC analysis revealed strong predictive ability for PROMIS-UE (AUC = 0.717, p < 0.05), moderate predictive ability for PROMIS-PI (AUC = 0.634, p < 0.05), and excellent predictive ability for PROMIS-D (AUC = 0.864, p < 0.05). Using 95% specificity, prognostic PROMIS domain cutoffs were generated that yielded 100% success in achievement of MCID, for all three PROMIS CAT domains Table 3A. Similarly, failure to achieve MCID was assessed through ROC analysis and prognostic cutoffs can be seen in Table 3B. For example, approximately 80% of patients achieve MCID at the 10-month postoperative time point, but 100% of patients that present with PROMIS-UE scores under 26.0 met MCID at the 10-month mark. Scatter plots of the change in PROMIS CAT domain by the preoperative t score are displayed in Figs. 1–3. These plots can be used to visualize patients achievement, or lack thereof, of MCID with respect to prognostic t score cutoffs. The multivariate model did not show any significant improvement in predicting achievement or failure to achieve MCID (p = .169).

### 3. Discussion

The results of this study demonstrate that patients undergoing primary reverse shoulder arthroplasty for rotator cuff tear arthropathy show significant improvements in PROMIS-UE, PROMIS-PI, and PROMIS-D at mean 10-months follow-up. Furthermore, clinical improvement can be measured using MCID values of 3.1, 2.7, and 4.6 for PROMIS-UE, PROMIS-PI, and PROMIS-D, respectively. Finally, patients may be counseled using prognostic cutoffs that suggest increased likelihood of MCID achievement: <26.0 for PROMIS-UE, >70.0 for PROMIS-PI, and >52.5 for PROMIS-D.

In recent years, multiple orthopedic articles have emerged reporting the validity and efficiency of PROMIS CAT use in the upper extremity [1,25]. Specifically, the responsiveness of these forms, or the ability to dynamically capture change over time, has been validated to track shoulder outcomes longitudinally [10]. Felicity et al showcased the ability of PROMIS-UE, PROMIS-PI, and PROMIS-D to dynamically change at multiple postoperative time points in patients undergoing rotator cuff repair, as late as 6-months postoperatively [10]. While our study only evaluated change at one time point, we also show a significant improvement in all three PROMIS CAT domains at the 10-month postoperative time point, suggesting the patient-centric impact of reverse shoulder arthroplasty can be tracked by the physician.

| Table 3A – Prognostic cutoffs and probability of MCID achievement. |
|--------------------------|--------------------------|--------------------------|--------------------------|
| Pre-Cutoff probability (%) | Prognostic cutoff | Post-Cutoff probability (%) | AUC |
| PROMIS-UE | 82.1% | <26.0 | 100.0% | 0.717 |
| PROMIS-PI | 85.2% | >70.0 | 100.0% | 0.634 |
| PROMIS-D | 55.6% | >52.5 | 100.0% | 0.864 |

Abbreviations: Minimal Clinically Important Difference (MCID); Area Under the Curve (AUC); Patient-Reported Outcomes Measurement Information System (PROMIS); Upper Extremity Physical Function (UE); Pain Interference (PI); Depression (D).

| Table 3B – Prognostic cutoffs and probability of failure to achieve MCID. |
|--------------------------|--------------------------|--------------------------|--------------------------|
| Pre-Cutoff probability (%) | Prognostic cutoff | Post-Cutoff probability (%) | AUC |
| PROMIS-UE | 17.9% | >36.6 | 57.1% | 0.717 |
| PROMIS-PI | 14.8% | <55.2 | 50.0% | 0.634 |
| PROMIS-D | 44.4% | <40.6 | 90.0% | 0.864 |

Abbreviations: Minimal Clinically Important Difference (MCID); Area Under the Curve (AUC); Patient-Reported Outcomes Measurement Information System (PROMIS); Upper Extremity Physical Function (UE); Pain Interference (PI); Depression (D).
Furthermore, we showed that these values change in accordance with clinical measures, such as range of motion. Ngan et al echoed this sentiment by showing strong correlation between PROMIS scores and functional workspace of patients after both total (n = 70) and reverse (34) shoulder arthroplasty intervention [19]. Combined with the efficiency and ease of administration outlined in the orthopedic literature [3,15,22], clinicians should consider adapting their patient-reported outcome collection to include these CAT domains in their practice.

**Fig. 1** – Scatter plots were generated to visualize participants change in upper extremity physical function in comparison to preoperative upper extremity physical function scores. Vertical dashed lines represent prognostic cutoffs for achieving MCID (26.0) and failing to achieve MCID (36.6). The horizontal axis delineates the MCID value (3.1). Each blue shaded area indicates participants who achieved (upper left) and failed to achieve (bottom right) MCID.

**Fig. 2** – Scatter plots were generated to visualize participants change in pain interference in comparison to preoperative pain interference scores. Vertical dashed lines represent prognostic cutoffs for achieving MCID (70.0) and failing to achieve MCID (55.2). The horizontal axis delineates the MCID value (2.7). Each blue shaded area indicates participants who achieved (bottom right) and failed to achieve (upper left) MCID.
MCID values in PROMIS CAT domains have also become increasingly reported in the orthopedic literature for a variety of operations [6,16,20]. Unfortunately, the evaluation of MCID in upper limb procedures is sparse when compared to the lower extremity. Chen et al recently made the first steps toward PROMIS CAT domain MCID analysis in shoulder arthroplasty patients [5]. With a cohort of 62 total shoulder arthroplasty patients, Chen et al identified MCID values of 4.0, 3.2, and 4.3 for PROMIS-UE, PROMIS-PI, and PROMIS-D, respectively. Furthermore, they introduced a multifactorial model (including age, sex, BMI, ASA class, and each PROMIS CAT domain) as a means for predicting MCID, which displayed superior predictive ability when compared to their univariate model. To our knowledge, the present study is the first to identify MCID values in a strictly reverse shoulder arthroplasty patient cohort. The present study assessed 73 patients with CPT code 24372, but delineated reverse shoulder arthroplasty patients for our analysis. Our proposed MCID values were marginally smaller (<1.0 difference), which may suggest more uniformity with preoperative upper limb functioning and pain interference. Although, without demographic analysis provided in the Chen et al article, we cannot compare the uniformity of these two cohorts. Furthermore, their predictive models all showed moderate ability to distinguish those achieving MCID, whereas PROMIS-UE showed strong ability and PROMIS-D showed excellent ability in our cohort. The strength of our univariate models may be able to explain the lack of added benefit when using a multivariate model in our cohort. These findings provide further evidence that PROMIS CAT domains can aid clinicians in patient selection and counseling when considering orthopedic intervention.

Using the prognostic cutoffs presented in Tables 3A-B, in accordance with clinical presentation and radiological exams, orthopedic surgeons are able to better counsel patients in their decision to undergo reverse shoulder arthroplasty. When a patient completes their baseline PROMIS CAT surveys, typically in a surgical consult appointment, the surgeon may discuss the impact of these patient-reported values on their probability of RSA providing significant benefit. If the hypothetical patient presents with a PROMIS-UE value of 25.3, a PROMIS-PI value of 61.1, and a PROMIS-D value of 40.2, the clinician may provide the patient with further context using Tables 3A-B. The PROMIS-UE score would meet the suggested prognostic cutoff for physical function gain (26.0), while both the PROMIS-PI and PROMIS-D scores would fall short of their prognostic cutoffs for significantly higher chances of mitigation of pain (70.0) and depression (52.5), respectively. More specifically, the patient would meet the prognostic cutoff that indicates a heightened probability to not experience significant mitigation of depression (40.6). Thus, the patient may expect to have a significantly higher chance to improve their physical function, an average chance of experiencing improvement in pain, and a low chance of experiencing improvement in their mental health. This additional level of insight provided may help both manage patient expectations as well as guide clinicians in altering postoperative procedures, such as physical therapy or pain management protocols.

Our study does present with notable limitations. Primarily, selection bias may have been present due to the variable nature of follow-up in patients undergoing reverse shoulder arthroplasty. Our follow-up period captured patients...
4. Conclusions

PROMIS domains can adequately measure reverse shoulder arthroplasty patients’ symptomatic states as late as 10-months postoperatively. Furthermore, preoperative baseline scores can serve as strong predictors of success in patients undergoing primary reverse shoulder arthroplasty and can be used to both counsel patients on surgery and to tailor postoperative protocols. However, this predictive ability should not be used as doctrine, rather as a facilitative tool for currently accepted orthopedic diagnostic measures.

Disclaimer

The 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


