J Shoulder Elbow Surg (2018) 27, 298–305

Journal of Shoulder and Elbow Surgery

www.elsevier.com/locate/ymse

BASIC SCIENCE

Quantifying success after total shoulder arthroplasty: the minimal clinically important difference

Ryan Simovitch, MD[a], Pierre-Henri Flurin, MD[b], Thomas Wright, MD[c], Joseph D. Zuckerman, MD[d,*], Christopher P. Roche, MSE, MBA[e]

[a]*Palm Beach Orthopaedic Institute, Palm Beach Gardens, FL, USA*
[b]*Department of Orthopedic Surgery, Bordeaux-Mérignac Clinic, Mérignac, France*
[c]*Department of Orthopedic Surgery, Department of Orthopedics, University of Florida, Gainesville, FL, USA*
[d]*Department of Orthopedic Surgery, NYU Langone Orthopaedic Hospital, New York, NY, USA*
[e]*Exactech, Inc., Gainesville, FL, USA*

Background: Knowledge of the minimal clinically important difference (MCID) for different shoulder outcome metrics and range of motion after total shoulder arthroplasty (TSA) can be useful to establish a minimum threshold of improvement that defines successful treatment. This study quantifies how MCID varies with different prosthesis types, patient age, gender, and length of follow-up after TSA.

Methods: A total of 466 anatomic TSA (aTSA) and reverse TSA (rTSA) with 2-year minimum follow-up were performed by 13 shoulder surgeons. The MCID for the American Shoulder and Elbow Surgeons, Constant, University of California Los Angeles Shoulder Rating Scale, Simple Shoulder Test, Shoulder Pain and Disability Index, global shoulder function, and visual analog scale for pain scores, as well as active abduction, forward flexion, and external rotation, were calculated for different prosthesis types and patient cohorts using an anchor-based method.

Results: The anchor-based MCID results were American Shoulder and Elbow Surgeons = 13.6 ± 2.3, Constant score = 5.7 ± 1.9, University of California Los Angeles Shoulder Rating Scale = 8.7 ± 0.6, Simple Shoulder Test score = 1.5 ± 0.3, Shoulder Pain and Disability Index score = 20.6 ± 2.6, global shoulder function = 1.4 ± 0.3, pain visual analog scale = 1.6 ± 0.3, active abduction = 7° ± 4°, active forward flexion = 12° ± 4°, and active external rotation = 3° ± 2°. Female gender and rTSA were associated with lower MCID values compared with male gender and aTSA patients.

Conclusion: The minimum improvement necessary for patients to achieve a result they believe is clinically meaningful after aTSA and rTSA is nominal and was achieved by at least 80% of the patients. Future endeavors should investigate the influence of different anchor questions on the MCID calculation.

Level of evidence: Basic Science Study; Validation of Outcome Instruments

Keywords: Minimal clinically important difference; MCID; anatomic total shoulder arthroplasty; reverse total shoulder arthroplasty; shoulder arthroplasty; shoulder replacement

Data acquisition and analysis was performed with approval from the Western Institutional Review Board (protocol # WIRB 20091701).

*Reprint requests: Joseph D. Zuckerman, MD, NYU Langone Orthopedic Hospital, Department of Orthopedic Surgery, 301 E 17th St, 14th Flr, New York, NY 10003, USA.

E-mail address: joseph.zuckerman@nyumc.org (J.D. Zuckerman).

1058-2746/$ - see front matter © 2017 Journal of Shoulder and Elbow Surgery Board of Trustees. All rights reserved.
https://doi.org/10.1016/j.jse.2017.09.013

Total shoulder arthroplasty (TSA) is a commonly accepted treatment for a variety of pathologies of the glenohumeral joint. A vast array of clinical and radiographic studies have demonstrated statistically significant improvements after both anatomic TSA (aTSA) and reverse TSA (rTSA).[5-7,9-11,13,21,23,24,27,28,30,33,36,37,39] Recently there has been interest in determining the minimal clinically important difference (MCID) after shoulder arthroplasty[8,32,34] to define the threshold of improvement that is clinically meaningful to a patient.

MCID was first defined by Jaeschke et al[14] in 1989 to quantify the smallest difference in a clinical outcome measure that a patient would perceive as a beneficial and meaningful change by a given treatment. The MCID ideally avoids identification of small changes in outcome measures that appear to be meaningful solely as the result of statistical significance, which is dependent on sample size and other study power–related variables.[18] Statistical significance can sometimes exist coincident with clinical irrelevance, as judged by a patient.

There are 3 accepted methods for determining MCID: the distribution, Delphi, and anchor methods. The distribution method relies on statistical testing, typically using a ratio of the standard deviation for a given metric. The Delphi method relies on repeated sampling of patients and experts to build MCID consensus. The anchor method evaluates clinical outcomes relative to a global question that can represent overall well-being or response to a surgical procedure or intervention.

Several different metrics are used to quantify outcomes after shoulder arthroplasty, including the Simple Shoulder Test (SST), University of California Los Angeles (UCLA) Shoulder Rating Scale, American Shoulder and Elbow Surgeon (ASES), Constant, Shoulder Pain and Disability Index (SPADI), and pain visual analog score (VAS) metrics.[5,6,11,12,21,27-29]

With outcomes being tied to reimbursement, MCID thresholds are increasingly relevant and necessary to define the standard threshold for successful treatment and to ensure judicious allocation of finite economic resources relative to the performance of shoulder arthroplasty.[1,4,26] To this end, an improved understanding is needed for how MCID varies across different study cohorts and patient populations after TSA. For example, MCID may be influenced by variables beyond the control of the surgeon such as patient age, gender, and length of postoperative follow-up.

The purpose of this study was to quantify the MCID for the SST, UCLA, ASES, Constant, and SPADI tests and for the pain VAS and global shoulder function scores. Finally, we quantified the effect of prosthesis type, patient age, gender, and length of follow-up on the MCID for each of the these outcome metrics to establish a minimum threshold for successful treatment.

Materials and methods

This was a retrospective outcome study focused on patients treated with aTSA and rTSA who were prospectively enrolled in a multicenter database. Between February 2001 and February 2015, data were collected on 2057 patients treated by 13 fellowship-trained orthopedic surgeons using primary aTSA or rTSA with a single platform shoulder system (Equinoxe; Exactech, Inc., Gainesville, FL, USA). Inclusion criteria included primary aTSA performed for osteoarthritis or rheumatoid arthritis and rTSA performed for cuff tear arthropathy or a combination of osteoarthritis and rotator cuff insufficiency. Patients with fracture diagnoses and revision cases were excluded. However, patients who were enrolled for a primary procedure and ultimately underwent revision for a complication were included in the analysis with the data recorded at the last visit before revision. Only patients with 2 years or more of clinical and radiographic follow-up were included.

The application of all inclusion and exclusion criteria resulted in 1856 patients (1098 women and 758 men) consisting of 911 aTSA and 945 rTSA shoulders. The average age was 69.6 ± 8.8 years (range, 31-93 years). The average body mass index was 28.8 ± 5.9 kg/m^2 (range, 17.0-48.1 kg/m^2). The average follow-up was 44.9 ± 23.8 months (range, 24-157 months). Differences in age, gender, body mass index, and length of follow-up between aTSA and rTSA patients are recorded in Table I.

Patients were evaluated preoperatively and at the latest follow-up using the SST, UCLA, ASES, Constant, and SPADI metrics. Pain was recorded on a VAS from 0 to 10 in increments of 1. A global shoulder function score was also recorded from 0 to 10 in increments of 1. The global shoulder function was assessed by asking a patient to rate his or her shoulder on a scale from 0 to 10, with 10 being most functional. Active abduction, forward flexion, and external rotation were also recorded. Patient interrogation, range of motion, and strength evaluation were performed by the procedural surgeon, a physical therapist, or research coordinator. Substantial effort was made to standardize the method of data collection and entry. Complications were recorded as well.

Radiographic analysis was conducted at the latest follow-up using anteroposterior, axillary lateral, and scapular Y x-ray imaging. Radiographs were evaluated for lucency around the humeral stem (aTSA and rTSA) and glenoid components (aTSA) according to the Gruen classification adapted to the humerus[20] and the Lazarus classification,[16] respectively. Scapular notching (rTSA) was recorded according to the Nerot classification.[36]

At the latest follow-up, a global anchor question asked each patient to rate his or her shoulder as "worse," "unchanged," "better," or "much better" relative to the preoperative condition. This anchor question was modeled after the anchor question used by Tashjian et al,[32] who evaluated shoulder arthroplasty by using a response to a treatment

Table I Comparison of demographics of the reverse and anatomic total shoulder arthroplasty cohorts

Demographics	aTSA cohort	rTSA cohort	*P* value
Gender			<.0001
Female	488	610	
Male	423	335	
Age, y	66.5 ± 9.1	72.5 ± 7.5	<.0001
BMI, mg/kg^2	29.5 ± 6.1	28.1 ± 5.6	<.0001
Follow-up, mo	49.7 ± 27.5	40.2 ± 18.6	<.0001

aTSA, anatomic total shoulder arthroplasty; *rTSA*, reverse total shoulder arthroplasty; *BMI*, body mass index.
Categoric data are shown as the number of patients and continuous data as the mean ± standard deviation.