J Shoulder Elbow Surg (2018) ■■, ■■–■■

ELSEVIER

Journal of Shoulder and Elbow Surgery

www.elsevier.com/locate/ymse

REVIEW ARTICLE

Establishing maximal medical improvement after anatomic total shoulder arthroplasty

Richard N. Puzzitiello, BS, Avinesh Agarwalla, BS, Joseph N. Liu, MD, Gregory L. Cvetanovich, MD, Anthony A. Romeo, MD, Brian Forsythe, MD*, Nikhil N. Verma, MD

Division of Sports Medicine, Midwest Orthopaedics at Rush, Rush University Medical Center, Chicago, IL, USA

Background: As a heightened emphasis continues to be placed on value-based health care, quality outcomes following orthopedic procedures must be properly defined. With knowledge of the time to maximal medical improvement following total shoulder arthroplasty (TSA), physician resources can be justly allocated to optimize value in ambulatory orthopedic care.

Materials and methods: A systematic review was conducted to identify studies reporting sequential follow-up at several time points, up to a minimum of 2 years after TSA. Assessment for clinically significant improvements between time intervals was made by using the minimal clinically important difference specific to each patient-reported outcome measure.

Results: We identified 13 studies that fit the criteria to be included in this review, amounting to 984 patients who underwent TSA. Clinically significant improvements in patient-reported outcome scores were appreciated up to 1 year following TSA, but no further clinical significance was seen from 1 year to 2 years. Objective physical examination measurements followed a similar trend, with clinically significant improvements in abduction occurring up to 1 year postoperatively. For both the subjective and objective outcomes, the majority of improvements occurred in the first 3 months after the procedure.

Conclusions: Following TSA, clinically significant improvements in patient-reported outcomes and objective clinical measurements are seen up to 1 year postoperatively but not beyond this time. This result is important for counseling patients and modifying their expectations prior to surgery as well as for establishing a time frame for maximized outcome evaluation to define the value received from TSA.

Level of evidence: Level IV; Systematic Review

Keywords: Total shoulder arthroplasty; maximal medical improvement; clinical significance; MCID; value; patient-reported outcomes; minimal clinically important difference

Anatomic total shoulder arthroplasty (TSA) continues to surge in the United States, with nearly 35,000 procedures performed in 2013[25] and an annual increase as high as 12.1%.[19] Recent projections have indicated that the demand for this procedure will increase by 755.4% by 2030.[19] As health care's share of the total US economic budget continues to grow,[12] a heightened emphasis is being placed on resource optimization and value-based health care. This trend involves a departure from previous volume-based models toward a focus on health outcomes achieved as a result of the provided care per dollar spent.[2] The issue that arises with outcome-based medicine is how to properly define and measure quality

*Reprint requests: Brian Forsythe, MD, Division of Sports Medicine, Midwest Orthopaedics at Rush, 1611 W Harrison St, Chicago, IL 60612, USA.
E-mail address: brian.forsythe@rushortho.com (B. Forsythe).

1058-2746/$ - see front matter © 2018 Journal of Shoulder and Elbow Surgery Board of Trustees. All rights reserved.
https://doi.org/10.1016/j.jse.2018.03.007

outcomes following procedures, as well as when to follow up with patients to capture maximal medical improvement (MMI) while avoiding "unnecessary" visits.

The outcome metrics that are most valued by patients in orthopedic surgery are improvements in quality of life, as measured by decreased pain and increased function.[18,24] These metrics are assessed by validated patient-reported outcome measures (PROMs), which produce summative scores of a patient's limitations, symptoms, and satisfaction.[24] When one is assessing recovery from orthopedic surgery, especially across large patient cohorts, statistically significant changes in these PROMs may not result in a detectable change in pain or function for the patient.[18] The minimal clinically important difference (MCID) of PROMs, or "the smallest difference in score . . . which patients perceive as beneficial,"[11] is a more valid assessment of meaningful clinical outcomes than statistical significance. Using the MCIDs for PROMs ensures that measurements of quality outcomes remain patient centered.[34]

TSA provides excellent pain relief and restoration of function in the short term to midterm[3,27] and has excellent long-term implant survivorship.[21] Current follow-up schedules after arthroplasty typically include patient visits empirically scheduled at several time points over a 2-year period.[29,32] These short-term clinic visits are typically purposed to assess patient recovery, while visits after 2 years are typically less frequent and exist to monitor for signs of late complications.[36] Once MMI is reached, clinic visits could potentially be deferred until later time points when relevant changes, such as glenoid loosening, are more likely to occur.[21] Reducing the number of inconsequential follow-up visits would improve health care efficiency and value while minimizing patient and provider burden. In addition, a time frame for outcome reporting should be established as value-based reimbursement schemes are evolving.

The purpose of this systematic review was to establish when maximal improvement occurs following TSA. We hypothesized that patients would continue to perceive improvements until 1 year after their operation but would detect no additional improvements between 1 and 2 years.

Materials and methods

Systematic review and data extraction

Two reviewers independently searched the MEDLINE database on October 17, 2017. The following search terms were used: "total shoulder arthroplasty" or "total shoulder replacement" in combination with "recovery," "outcome," or "clinical results." Studies were included if they reported clinical outcomes of anatomic TSA, using either a stemmed or stemless humeral implant, for the indication of glenohumeral osteoarthritis (OA), with outcomes reported for at least 2 separate postoperative time points with a minimum of 2 years' follow-up. Articles were excluded if a TSA was not performed, if outcomes were not listed in numerical form, if outcomes at 2 years were not reported, or if outcomes were reported at only 1 postoperative time (Fig. 1). Full-text articles were evaluated if inclusion of the study was considered or if there was uncertainty about a study. If a study's methods seemed to meet the inclusion criteria but there were insufficient data reported, the corresponding author was contacted for the data. We allowed 4 weeks for the corresponding author to respond; otherwise, the study was excluded. The citations of each included study were also independently reviewed for articles that may have been missed on the initial search. If disagreement existed regarding inclusion of a study, the reviewers discussed it to reach the final determination.

The following PROMs were extracted from the articles fitting the inclusion criteria: Western Ontario Osteoarthritis of the Shoulder index; American Shoulder and Elbow Surgeons (ASES) score; Short Form 12 (SF-12) Physical and Mental Health Summary Scale scores; Simple Shoulder Test (SST) score; Single Assessment Numeric Evaluation score; Penn shoulder score; University of California, Los Angeles shoulder score; visual analog scale (VAS) pain and function score; absolute Constant-Murley score (ACMS); relative Constant-Murley score (RCMS); Quick Disabilities of the Arm, Shoulder and Hand (QuickDASH) score; and Shoulder Pain and Disability Index. Clinical examination data for active range of motion (ROM) and strength were also extracted when reported.

The MEDLINE database was searched for articles elucidating the MCID after TSA for OA for each PROM. The following search terms were used: "MCID" or "minimally clinically important difference" in combination with "total shoulder replacement" or "total shoulder arthroplasty."

Data analysis

Data were pooled and analyzed separately for each outcome score using the techniques previously described by Zuke et al.[39] The weighted means for each study that reported outcomes at a given time point were pooled, and the pooled standard deviation was then calculated. If a single study reported a particular PROM at a given time point and if that study did not report a mean score with a standard deviation, then that PROM was not analyzed at that time point. The pooled weighted means were compared at the following intervals: preoperatively to 3 months (or 6 months if data at 3 months were not available), 3 months to 6 months, 3 months to 1 year, 6 months to 1 year, 6 months to 2 years, and 1 year to 2 years. Nonconsecutive time points were analyzed to help further elucidate the point of MMI. A clinically significant improvement between time points was established if an improvement in an outcome score significantly exceeded the previously established MCID for the specific outcome measure ($P < .05$). MCID was first described by Jaeschke et al[11] as "the smallest difference in score…which patients perceive as beneficial." To remain consistent with this definition of MCID, if multiple MCIDs were previously reported for an individual PROM, the smallest MCID was used for analysis. This was done rather than finding an average among the scores because a discernible change in pain or function was noted by patients at the lowest MCID, which justifies a minimal clinically important improvement in outcome score. In addition, this errs on the side of more frequent visits owing to increased sensitivity of detecting change. The following MCIDs were used for analysis: ASES score, 6.3[37]; SST score, 1.8[30]; ACMS, 5.7[30]; Penn shoulder score, 11.4[13]; VAS pain score, 1.4[31]; Shoulder Pain and Disability Index, 20.6[30]; SF-12 Mental Health Summary Scale score, 5.7[38]; and SF-12 Physical Health Summary Scale score, 5.4.[38] Clinical significance could not